

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
31 January 2002 (31.01.2002)

PCT

(10) International Publication Number
WO 02/08245 A2

(51) International Patent Classification⁷: C07K 1/00

(21) International Application Number: PCT/US01/22194

(22) International Filing Date: 13 July 2001 (13.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/218,113 13 July 2000 (13.07.2000) US
PCT/US00/26821

29 September 2000 (29.09.2000) US

60/279,302 28 March 2001 (28.03.2001) US

09/839,577 20 April 2001 (20.04.2001) US

(71) Applicant (for all designated States except US): THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10555 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): HAHN, Klaus, M. [US/US]; 10258 Saunders Drive, San Diego, CA 92131 (US). TOUTCHKINE, Alexei [US/US]; 1647 Grand Avenue, Apartment 3, San Diego, CA 92109 (US). MUTHYALA, Rajeev [IN/US]; 1611 Valley Road, Apartment B2, Champaign, IL 61820 (US). KRAYNOV, Vadim [US/US]; 7873 Avenida Navidad, Apartment 267, San Diego, CA 92131 (US). BARK, Steven, J.

[US/US]; 10730 Carillon Court, San Diego, CA 92131 (US). BURTON, Dennis, R. [US/US]; 6044 Beaumont Avenue, LaJolla, CA 92037 (US). CHAMBERLAIN, Chester [US/US]; 10925 Decatur Road, San Diego, CA 92126 (US).

(74) Agent: VIKSNINS, Ann, S.; Schwegman, Lundberg, Woessner & Kluth, P.O. Box 2938, Minneapolis, MN 55402 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

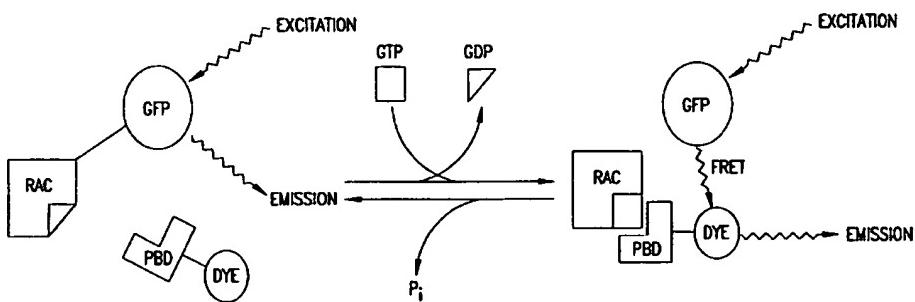
— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: LABELED PEPTIDES, PROTEINS AND ANTIBODIES AND PROCESSES AND INTERMEDIATES USEFUL FOR THEIR PREPARATION



WO 02/08245 A2



(57) Abstract: The invention provides peptide synthons having protected functional groups for attachment of desired moieties (e.g. functional molecules or probes). Also provided are peptide conjugates prepared from such synthons, and synthon and conjugate preparation methods including procedures for identifying optimum probe attachment sites. Biosensors are provided having functional molecules that can locate and bind to specific biomolecules within living cells. Biosensors can detect chemical and physiological changes in those biomolecules as living cells are moving, metabolizing and reacting to its environment. Methods are included for detecting GTP activation of a Pho GTPase protein using polypeptide biosensors. When the biosensor binds GTP-activated Rho GTPase protein, an environmentally sensitive dye emits a signal of a different lifetime, intensity or wavelength than when not bound. New fluorophores whose fluorescence responds to environmental changes are also provided that have improved detection and attachment properties, and that can be used in living cells, or in vitro.

WO 02/08245 A2



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

**Labeled Peptides, Proteins and Antibodies and Processes and Intermediates
Useful for Their Preparation**

Government Funding

5 The invention described herein was made with United States Government support under Grant Number MCB-9812248 awarded by the National Science Foundation, and under Grant Numbers CA58689, AG15430 and GM 57464 awarded by the National Institutes of Health. The United States Government has certain rights in this invention.

10

Background of the Invention

Modified peptides and proteins are valuable biophysical tools for studying biological processes, both *in vitro* and *in vivo*. They are also useful in assays to identify new drugs and therapeutic agents. In particular, quantitative live cell imaging using fluorescent proteins and peptides is revolutionizing the study of cell biology. An exciting recent development within this field has been the construction of peptide and protein biosensors exhibiting altered fluorescence properties in response to changes in their environment, oligomeric state, conformation upon ligand binding, structure, or direct ligand binding.

15 Appropriately labeled fluorescent biomolecules allow spatial and temporal detection of biochemical reactions inside living cells. See for example Giuliano, K.A., et al., *Annu. Rev. Biophys. Biomol. Struct.* 1995, 24:405-434; Day, R.N., *Mol. Endocrinol.* 1998, 12:1410-9; Adams, S.R., et al., *Nature* 1991, 349:694; Miyawaski, A., et al., *Nature* 1997, 388:882-7; Hahn, K., et al., *Nature* 1992, 25: 359:736; Hahn, K.M., et al., *J. Biol. Chem.* 1990, 265:20335; and Richieri, G.V., et al., *Mol. Cell. Biochem.* 1999, 192:87-94.

20

Procedures for site-specific modification of polypeptides have been described, including: chemically selective labeling in solution (Brinkley, M. *Bioconjugate Chemistry* 1992, 3:2-13) and on resin bound peptides (Hackeng, T., et al., *J. Biol. Chem.* Submitted); introduction of ketone amino acids through synthetic procedures (Rose, K. *J. Am. Chem. Soc.* 1994, 116:30-33; King, T.P., et al., *Biochemistry*, 1986, 25:5774-5779; Rose, K., et al., *Bioconjugate Chem.* 1996, 7:552-556; Marcaurelle, L.A., Bertozzi, C.R. *Tet. Lett.* 1998, 39:7279-

7282; and Wahl, F., Mutter, M. *Tett. Lett.* 1996, 37:6861-6864); and molecular biology techniques (Cornish, V.W., et al., *J. Am. Chem. Soc.* 1996, 118:8150). For example, green fluorescence protein (GFP) is a fluorescent protein that has been fused to other proteins using molecular biology techniques and has been 5 used to visualize intracellular proteins (see, e.g., Katz et al., *BioTechniques* 25: 298-304 (1998)).

While each of these methods has utility for producing a particular class of biosensor or labeled polypeptide, all have limitations that restrict their general use. Labeling of natural amino acid side-chains in solution is often impractical 10 because of the existence of many other competing nucleophiles. Additionally, the use of unnatural amino acids, such as those bearing ketones for selective labeling, requires the synthesis of dye constructs or amino acids that are difficult to make and are not available commercially. While a variety of proteins have been fused to GFP, some GFP-labeled proteins fail to fluoresce. Mutational 15 analysis indicates that the structure of GFP is extremely sensitive to molecular or biochemical modifications. Moreover, GFP is extremely large (238 amino acids). Hence, fusion of GFP to other proteins can alter the function of GFP or even the function of the protein to which it is fused. Moreover, GFP does not have the diverse capabilities of smaller, synthetic molecules, some of which 20 provide a variety of fluorescence wavelengths, or are capable of reporting on protein conformation, or can photo-crosslink or act as NME or EPR probes. Accordingly, new fluorescent molecules with more diverse properties and new attachment methods are needed.

Currently, the major obstacles to the development of fluorescent 25 biosensors and labeled polypeptides remain: (1) The difficulty in site-specific placement of the dye in the polypeptide and (2) determining exactly which site is optimal for dye placement (Giuliano, K.A., et al., *Annu. Rev. Biophys. Biomol. Struct.* 1995, 24:405-434). Solvent-sensitive dyes and other biophysical probes must be placed precisely for optimal response to changes in protein structure 30 without interference with biological activity. Also, the need for site-specific incorporation of two dyes without impairment of biological activity has proven a serious limitation for utilization of fluorescence resonance energy transfer

- (FRET) within a single protein. Total chemical synthesis of proteins provides a potential solution to these problems (Wilken, J., Kent, S.B.H. *Curr. Op. Biotechnology*. 1998, 9:412; Kent, S.B.H. *Ann. Rev. Biochem.* 1988, 57, 957-989; Dawson, P.E., et al., *Science* 1994, 266:776-779; Muir, T.W., et al., *Proc. Natl. Acad. Sci.* 1998, 95:6705-6710; and Cotton, G.J., et al., *J. Am. Chem. Soc.* 1999, 121:1100-1101). However, many biophysical probes suitable for fluorescent biosensors or other purposes are not stable to the various conditions used for peptide synthesis, and site-specific incorporation after synthesis has been difficult to achieve.
- Moreover, labeling with hydrophobic dyes such as thionine or methylene blue can be problematic because these dyes autoaggregate in aqueous solution at high concentration. See, *J. Am. Chem. Soc.* 63, 69 (1941). These aggregates cause a change in the absorption spectrum and a reduction in the fluorescence of the dyes. Cyanines and merocyanines are also thought to aggregate, causing a quenching of fluorescence (*J. Phys. Chem.* 69, 1894 (1965)). Such aggregation interferes with conjugation of these fluorescent dyes to other molecules such as proteins. Moreover aggregation by cyanines and merocyanines can be exacerbated after the dyes are conjugated. For example, Waggoner et al. have observed an aggregation phenomenon following the conjugation of cyanin isothiocyanate with an antibody (*Cytometry* 10, 11-19 (1989)). Fluorescence of a conjugate between a cyanin fluorescent dye and an anti-HCG antibody (molar ratio=1.7) named CY5.18 is quenched in comparison with that of the free cyanin (see U.S. Pat. No. 5,268,486 and Anal. Biochem. 217, 197-204 (1994)). Also, while cyanines are generally stable, inexpensive, simple to conjugate to other molecules and of a suitable size for the recognition of small molecules, they do not change their fluorescence in response to environmental factors, such as solvent polarity. New dyes that eliminate these problems are needed.
- Thus, there is currently a need for new fluorescent dyes and peptide synthons having protected functional groups that can be selectively modified to incorporate one or more functional molecules (e.g. a fluorescent label) following peptide synthesis. There is also a need for proteins and antibodies with biophysical probes attached to precise locations, and for simple, non-destructive

methods of making such labeled proteins and antibodies. Simpler methods for using these labeled peptides, proteins and antibodies *in vivo* as biosensors are also needed.

5

Summary of the Invention

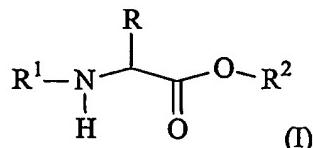
The present invention provides a highly efficient method for the site-specific attachment of biophysical probes or other molecules to unprotected peptides following chemical synthesis. The methodology utilizes amino acids having one or more protected aminoxy groups, which can be incorporated during solid-phase peptide synthesis or which can be combined with recombinant peptides through post expression steps. It has been discovered that the protected aminoxy group can be unmasked following peptide synthesis, and reacted with an electrophilic reagent to provide a modified (e.g. a labeled) peptide. The aminoxy group reacts selectively with electrophiles (e.g. an activated carboxylic ester such as an N-hydroxy-succinimide ester) in the presence of other nucleophilic groups including cysteine, lysine and amino groups.

Thus, selective peptide modification (e.g. labeling) can be accomplished after synthesis using commercially available and/or chemically sensitive

molecules (e.g. probes). The methodology is compatible with the synthesis of C^α-thioester containing peptides and amide-forming ligations, required steps for the synthesis of proteins by either total chemical synthesis or expressed protein ligation. An aminoxy containing amino acid can be introduced into different sites by parallel peptide synthesis to generate a polypeptide analogue family with each member possessing a single specifically-labeled site. The parallel synthesis enables the development of optimized biosensors or other modified polypeptides through combinatorial screening of different attachment sites for maximal response and minimal perturbation of desired biological activity.

Thus, a simple and efficient methodology for site-specific modification (e.g. labeling) of peptides after synthesis has been developed that provides high yield, selectivity, and compatibility with both solid-phase peptide synthesis and C^α-thioester peptide recombinant synthesis.

Accordingly, the invention provides a synthetic intermediate (i.e. a synthon) useful for preparing modified peptides, which is a compound of formula (I):



wherein:

5

R^1 is hydrogen or an amino protecting group;

R^2 is hydrogen or a carboxy protecting group;

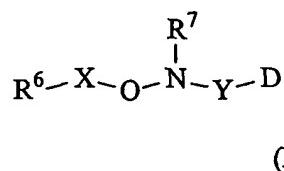
R is an organic radical comprising one or more aminoxy groups.

- Peptides including one or more aminoxy groups are also useful
- 10 synthetic intermediates that can be modified to provide related peptides having altered biological, chemical, or physical properties, such as, for example, a peptide linked to a fluorescent label. Accordingly, the invention also provides a peptide having one or more (e.g. 1, 2, 3, or 4) aminoxy groups; provided the peptide is not glutathione. The invention also provides a peptide having one or
- 15 more (e.g. 1, 2, 3, or 4) secondary aminoxy groups.

The invention generally provides intermediates and methods that allow for site-specific modification of peptides after synthesis. Accordingly, functional molecules can be selectively linked to a peptide to provide a peptide conjugate having altered biological, chemical, or physical properties. For

20 example, functional molecules (e.g. biophysical probes, peptides, polynucleotides, and therapeutic agents) can be linked to a peptide to provide a peptide conjugate having differing and useful properties.

Thus, the invention also provides a compound of formula (III):



25 wherein:

R⁶ is a peptide, polypeptide or antibody;
X is a direct bond or a linking group;
R⁷ is hydrogen, (C₁-C₆)alkyl, an amino protecting group, or a radical comprising one or more aminoxy groups;
5 Y is a direct bond or a linking group; and
D is a functional molecule.

A functional molecule can be any label, dye, pharmaceutical, toxin. Preferably the functional molecule is a biophysical probe, such as a fluorescent group that can be used for FRET studies or other studies involving fluorescent signals, such
10 as excimer pair formation.

Processes for preparing synthons of the invention as well as the polypeptide, antibody and protein conjugates of the invention are provided as further embodiments of the invention and are illustrated by the procedures in the Examples below.

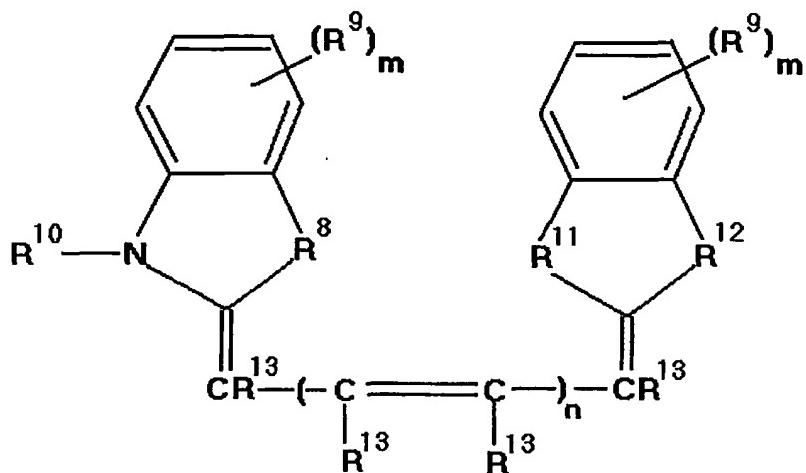
15 Thus, the invention also provides a method for preparing a peptide conjugate comprising a peptide and a functional molecule, comprising reacting a peptide having one or more aminoxy groups with a corresponding functional molecule having an electrophilic moiety, to provide the peptide conjugate.

The present invention further provides environment-sensing dyes that can
20 be readily conjugated to proteins and other molecules without the problems of aggregation, fluorescence quenching and the like. The present dyes strongly fluoresce and can fluoresce in an environmentally-sensitive manner suitable for use in living cells. Unlike cyanine dyes, the environmental sensitivity of these dyes can be used to form biosensors that can report many aspects of protein
25 behavior, or the behavior of other molecules. Protein behaviors including conformational change, phosphorylation state, ligand interaction, protein-protein binding and various post-translational modifications affect the distribution of charged and hydrophobic residues can be reported by the present dyes by changes in their fluorescence.

30 The present invention overcomes the disadvantages of the available environmentally-sensitive fluorescent dyes. The present fluorescent probes exhibit high fluorescence levels before and after conjugation to other molecules,

including peptides, proteins and antibodies, and changes in fluorescence suitable for many purposes, including *in vivo* and *in vitro* assays of protein behavior.

The present invention provides new fluorescent dyes that can be used in any manner chosen by one of skill in the art. The dyes can be linked to any useful molecule known to one of skill in the art using any available procedure. In one embodiment the fluorescent dyes are linked to peptides, polypeptides or antibodies using the methods provided herein. These dyes have the following structure (IV).



10

wherein:

each m is separately an integer ranging from 1-3;

n is an integer ranging from 0 to 5;

R⁸, R¹¹ and R¹² are separately CO, SO₂, C=C(CN)₂, S, O or

15 C(CH₃)₂;

each R¹³ is alkyl, branched alkyl or heterocyclic ring derivatized with charged groups to enhance water solubility and enhance photostability;

20 R⁹ and R¹⁰ are chains carrying charged groups to enhance water solubility (i.e. sulfonate, amide, ether) and/or chains bearing reactive groups for conjugation to other molecules. The reactive group is a functional group that is chemically reactive (or that can be made chemically reactive) with functional groups typically found in biological materials, or functional groups that can be

readily converted to chemically reactive derivatives using methods well known in the art. In one embodiment of the invention, the charged and reactive groups are separately haloacetamide (-NH-(C=O)-CH₂-X), where X is Cl, Br or I. Alternatively, the charged and reactive groups are separately amine, maleimide, 5 isocyanato (-N=C=O), isothiocyanato(-N=C=S), acyl halide, succinimidyl ester, or sulfosuccinimidyl ester. In another embodiment, the charged and reactive groups are carboxylic acid (COOH), or derivatives of a carboxylic acid. An appropriate derivative of a carboxylic acid includes an alkali or alkaline earth metal salt of carboxylic acid. Alternatively, the charged and reactive groups are 10 reactive derivatives of a carboxylic acid (-COORx), where the reactive group Rx is one that activates the carbonyl group of -COORx toward nucleophilic displacement. In particular, Rx is any group that activates the carbonyl towards nucleophilic displacement without being incorporated into the final displacement product. Examples of COORx : ester of phenol or naphtol that is further 15 substituted by at least one strong electron withdrawing group, or carboxylic acid activated by carbodiimide, or acyl chloride, or succinimidyl or sulfosuccinimidyl ester. Additional charged and reactive groups include, among others, sulfonyl halides, sulfonyl azides, alcohols, thiols, semicarbazides, hydrazines or hydroxylamines.

20 The invention still further provides a method of identifying an optimal position for placement of a functional molecule on a peptide having a peptide backbone and a known activity, which includes making a series of peptide conjugates, each peptide conjugate having the same amino acid sequence and the same functional molecule, wherein the functional molecule is linked at a 25 different location along the backbone of every peptide conjugate in the series, and observing which functional molecule location does not substantially interfere with the known activity of the peptide.

The invention also provides a method of identifying an optimal position for placement of a functional molecule in a protein having a known 30 activity and an identified peptide segment for attachment of the functional molecule, which includes making a series of peptide conjugates, each peptide conjugate having the amino acid sequence of the identified peptide segment and

the same functional molecule, wherein the functional molecule is linked at a different location along the backbone of every peptide conjugate in the series; replacing the identified peptide segment in each protein of a series of said proteins with a peptide conjugate selected from the series of peptide conjugates

5 to create a series of protein conjugates each having the functional molecule at a different location; and observing which functional molecule location does not substantially interfere with the known activity of the protein.

The invention further provides a method of identifying an optimal position for placement of an environmentally-sensitive functional molecule on a

10 peptide biosensor having a backbone, which includes making a series of peptide conjugates, each peptide conjugate having the same amino acid sequence and the same functional molecule, wherein the functional molecule is at a different location along the backbone of every peptide conjugate in the series, and observing which functional molecule location provides the strongest signal

15 change in response to an environmental change in the peptide conjugate. The signal change can be any observable change in a signal, for example, the change can be a change in fluorescence emission intensity, fluorescence duration or fluorescence emission wavelength. The environmental change in the peptide biosensor can be, for example, an interaction with a target molecule.

20 The invention still further provides a method of identifying an optimal position for placement of an environmentally-sensitive functional molecule in a protein having a known activity and an identified peptide segment for attachment of the functional molecule, which includes making a series of peptide conjugates, each peptide conjugate having the amino acid sequence of the

25 identified peptide segment and the same environmentally-sensitive functional molecule, wherein the environmentally-sensitive functional molecule is linked at a different location along the backbone of every peptide conjugate in the series; replacing the identified peptide segment in each protein of a series of said proteins with a peptide conjugate selected from the series of peptide conjugates

30 to create a series of protein conjugates, each having the environmentally-sensitive functional molecule at a different location; and observing which

functional molecule location provides the strongest signal change in response to an environmental change in the protein conjugate.

- The invention also provides a method for detecting GTP activation of a Rho GTPase protein, which includes contacting a polypeptide biosensor with a
- 5 test substance, wherein said polypeptide biosensor comprises a polypeptide capable of binding a GTP-activated Rho GTPase protein, and wherein said polypeptide is operatively linked to an environmentally sensitive fluorescent dye; and observing fluorescence emissions from the polypeptide biosensor at a wavelength emitted by said fluorescent acceptor dye; wherein the
- 10 environmentally sensitive fluorescent dye will emit light of a different intensity or a different wavelength when the polypeptide biosensor is bound to the GTP-activated Rho GTPase protein than when the polypeptide biosensor is not bound.

The invention further provides a method of detecting the location of a cellular protein within a living cell that includes providing the living cell with a

15 biosensor capable of binding to a tag on the cellular target protein; and detecting the location of a functional molecule on the biosensor within the living cell.

The tag on the cellular target protein can be a peptide segment that has been fused to the cellular protein. In one embodiment, the tag is a peptide which includes SEQ ID NO:16 and that can bind to a biosensor having a peptide

20 segment with SEQ ID NO:15. The biosensor can include a peptide-conjugate of the invention. Any of the functional molecules, pharmaceuticals, toxins, labels, dyes or compounds can also be present on the biosensor. Moreover, any cellular protein can be detected using this method, for example, calmodulin, Rho GTPase, rac, cdc42, mitogen-activated protein kinase, Erk1, Erk2, Erk3, Erk4,

25 IgE receptor (F_eRI), actin, α -actinin, myosin, or a major histocompatibility protein. The methods provided herein can detect and identify the cellular location of proteins and cellular proteins that have never been successful labeled and observed *in vivo*.

30

Brief Description of the Figures

In the following detailed description of example embodiments of the invention, reference is made to the accompanying drawings which form a part

hereof, and which is shown by way of illustration only, specific embodiments in which the invention may be practiced. It is to be understood that other embodiments may be utilized and structural changes may be made without departing from the scope of the present invention.

5 FIG. 1 illustrates a general strategy for site-specific labeling of polypeptides. The protected aminoxy group is incorporated during solid-phase peptide synthesis (synthesis on a thioester-linker resin is shown); cleavage from the resin generates a peptide possessing unprotected sidechains, an aminoxy group and a C-terminal thioester; and ligation and subsequent site-specific
10 labeling produces the full-length peptide with a functional molecule attached at the aminoxy nitrogen.

FIG. 2 illustrates the synthesis of PA-test and SA-test peptides.

FIG. 3 shows HPLC analysis of purified SA-test peptide (top) and crude reaction products from optimized labeling conditions (bottom).

15 FIG. 4 illustrates the synthesis of a protected intermediate (4) of the invention.

FIG. 5 illustrates one method for using biosensors according to the present invention. In this example, the activation of Rac1 ("RAC") by GTP was observed. A fragment of p21-Activated kinase (PAK) capable of binding to
20 Rac1 (termed "PBD") was used as a biosensor because PAK will only bind to Rac1 when Rac1 is activated by GTP. A Rac1-Green Fluorescent Protein fusion protein (shown as a square named "RAC" attached to a circle named "GFP") was made and a cell line expressing this fusion protein was generated. Cells expressing RAC-GFP were injected with PBD labeled with Alexa-546 dye
25 ("Dye"). The PBD biosensor binds selectively to GTP-RAC-GFP, but not to GDP-RAC-GFP. Upon binding to GTP-RAC-GFP, the Alexa on the labeled fragment undergoes fluorescence resonance energy transfer (FRET) as the Alexa and GFP fluorophores are brought close together. This FRET can be measured within a living cell or in vitro to map the distribution, localization and level of
30 Rac-GTP activation. FRET produces a fluorescence signal which is distinct from a GFP fluorescence signal because energy is transferred from the excited GFP fluorophore to the nearby Alexa dye (J.R. Lakowicz, *Principles of*

Fluorescence Spectroscopy (Plenum Press, New York, 1983), pp.305-341)). By imaging the cell with different wavelengths, both the distribution of Rac and Rac activation can be studied in the same cell. GFP excitation and emission are used for overall Rac distribution, while GFP excitation and Alexa emission are used 5 for FRET.

FIG. 6 illustrates one method of using an environmentally sensitive fluorescent dye in the present methods so that changes in naturally existing proteins can be detected and observed *in vivo* or *in vitro*.

Figure 6a depicts FRET between the Rac1-Green Fluorescent Protein 10 (GFP-Rac1) fusion and the p21-activated kinase biosensor (PBD) labeled with Alexa-546 dye as shown in Figure 5. Inactive Rac1 is depicted as a larger gray circle with Green Fluorescent Protein (smaller circle) attached. Upon activation by GTP, Rac1 undergoes a structural change depicted as a gray circle changing to a half-rounded gray rectangle. Unbound PBD is depicted as a black L-shape 15 with an attached Alexa-546 dye (open circle). Before Rac1 is activated, PBD cannot bind and the Alexa-546 cannot undergo FRET. However, after Rac1 activation, the Rac1 assumes a conformation that permits PBD binding. Such binding juxtaposes the Green Fluorescence Protein and the Alexa-546, which produces FRET.

20 Figure 6b illustrates how an environmentally sensitive fluorescent dye eliminates the need to create a fusion protein like the GFP-Rac1 protein depicted in Figure 6a. In Figure 6b, a natural, unmodified protein is depicted as a gray oval. The protein changes conformation upon activation by GTP, depicted as the transition to a half-rounded gray rectangle. When the protein is in the activated 25 state, a polypeptide-biosensor that binds only to the activated state (black L-shape), with an attached environmentally sensitive dye (open circle), can bind. Upon binding, the environmentally sensitive dye will emit light of a different wavelength, duration or intensity (filled circle) than before binding. Use of this type of environmentally sensitive dye is further illustrated in the Figures to 30 follow.

FIG. 7 illustrates what conditions will optimally provide FRET between GFP-Rac and Alexa-PBD *in vitro*. Figure 7A shows fluorescence emission from

solutions containing a fixed level of GFP-Rac bound to GTP γ S at different concentrations of Alexa-PBD (PBD labeled with Alexa-546). Light at 480 nm was selectively used for GFP excitation, and direct (non-FRET) excitation of Alexa was subtracted from these spectra. In the absence of Alexa-PBD, the 5 emission from GFP (peak at 508 nm) is maximal and no Alexa emission (peak at 568 nm) is seen. As the concentration of Alexa-PBD is increased, binding of Alexa-PBD to Rac-GFP leads to FRET, producing increasing emission at 568 nm and a decrease at 508 nm. Figure 7B shows the variation of the 568/508 nm emission ratios with changes in the level of GTP (solid circles) or GDP (solid 10 squares). All data points were the average of three independent experiments.

FIG. 8 illustrates how GFP-Rac expression levels and levels of intracellular Alexa-PBD (as observed by fluorescence) correlate with changes in normal cell behavior produced by these proteins. Figure 8A shows what levels of GFP-Rac expression, as measured by log GFP intensity per cell area, were 15 correlated with ruffling. Cells with different expression levels of either wild type or a constitutively active Q6 IL mutant of GFP-Rac were scored for ruffling. Each point represents an individual cell, placed in the higher (Ruffling) or lower (Nonruffling) row depending on whether ruffling was induced. As illustrated, there is a level of GFP intensity below which ruffling was 20 consistently not induced by expression of wild type GFP-Rac. Only cells with Rac expression levels below 250 on this scale were used in biological experiments. The validity of this approach was supported by scoring of constitutively active Rac, which showed ruffle induction at much lower levels of expression.

Figure 8B shows which levels of intracellular Alexa-PBD would perturb 25 normal serum-induced ruffling, as observed by Alexa intensity per cell area. Cells were scored as in Figure 8A, at different levels of Alexa-PBD fluorescence. Based on this experiment, only cells with Alexa-PBD concentrations below 400 intensity units on this scale were used in biological 30 experiments. These studies demonstrated that FRET could still be readily detected at appropriately low levels of introduced protein.

FIG. 9 illustrates the dynamics of Rac activation during growth factor stimulation of quiescent cells. Figure 9A provides photomicrographs showing Rac localization (GFP-Rac) and Rac activation (FRET) before stimulation of quiescent Swiss 3T3 fibroblasts. Figure 9B provides photomicrographs of the 5 same Swiss 3T3 fibroblasts three minutes after addition of serum. Warmer or lighter colors correspond to higher intensity values. The cells showed accumulation of Rac at and around the nucleus before stimulation (GFP-Rac image). Most of the nuclear GFP-Rac was associated with the nuclear envelope. Serum or PDGF addition generated multiple moving ruffles that showed FRET, 10 while no FRET was seen at the nucleus before or after stimulation. Of thirty-five cells stimulated with either serum or PDGF, thirty-one began ruffling within 15 minutes. FRET was seen in the ruffles of all but one of the ruffling cells.

Figures 9C and 9D demonstrate that simple localization of Alexa-PBD is inferior to FRET in quantifying and localizing Rac-GTP binding (Bar = 8 μ m).
15 The ruffle in Figure 9B is shown in close-up in Figure 9C, visualized using FRET. PBD localization in the same region is visualized using Alexa fluorescence in Figure 9D, with scaling optimized for detection of the ruffle. Without prior knowledge of the ruffle's location, this localization would have been difficult to discern. The high background due to unbound PBD cannot be
20 eliminated in Figure 9D and binding to other target proteins is not eliminated as it is in the highly specific FRET signal.

In each of the GFP-Rac images above, intensities range between 300-1100. The image of FRET before serum addition was scaled to demonstrate the low levels of FRET, with values ranging between 0 and 15. In the image of
25 FRET after stimulation, the ruffle contains the highest values of 40 to 65. Nuclear FRET was not seen in any of the cells examined.

FIG. 10 illustrates Rac activation in motile Swiss 3T3 fibroblast cells. This figure shows two examples of where a Rac 1 activation gradient is formed, in confluent monolayer cells and in "wound healing" cells. High Rac1 activation
30 occurred at the leading edge of motility, particularly in wound healing cells. In these experiments high levels of Rac-GTP were frequently seen in the juxtanuclear region of the cell (Fig. 10A). The strong correlation of this gradient

with the direction of movement indicates that activated Rac is spatially organized in polarized cells to help guide or propagate movement. Comparison of the GFP (Fig. 10A) and FRET (Fig. 10B) images shows that the distribution of activated Rac does not parallel that of Rac 1 itself. FRET intensities (Fig. 5 10B) are 0-18 (top image) and 0-32 (bottom image). In the GFP images (Fig. 10A), intensities range between 98-700 (top image) and 100-1100 (bottom image).

FIG. 11 provides the structure of one of the present fluorescent dyes and the spectrum of fluorescence emission for that dye in water, methanol and 10 butanol. As illustrated, the fluorescent emission of this dye increases with increasing solvent hydrophobicity. This figure illustrates the environmental-sensitivity of this dye.

FIG. 12 provides the structure of another fluorescent dye of the present invention and shows its spectrum in aqueous solution, compared to similar dyes 15 lacking the groups designed to prevent aggregation. In the curve from each dye, the peak to the right is the unconjugated, highly fluorescent form of the dye, while that to the left is the weakly fluorescent form. The curve furthest to the right is the dye containing groups to prevent aggregation. As illustrated, the chosen groups help prevent aggregation of the dye.

20 FIG. 13 provides one method for synthesizing a dye of the present invention. Conversion of compound 1 to an amine 2 followed by protection of the amine provides compound 3, which can be alkylated to give compound 4. Reaction of compound 4 with compound 9 provides compound 5, which can be 25 deprotected to provide amine 6. Alkylation or acylation of amine 6 with a chain carrying a charged group, or with a chain bearing a reactive group for conjugation to another molecule, or with another molecule directly provides a dye of the invention 7. Intermediate compound 9 can be prepared by condensation of compound 8 with the requisite aldehyde, under conditions that are known in the art

30 FIG. 14 provides a three-dimensional image of the CRIB domain (“CBD”) of the Wiscott-Aldrich syndrome protein (WASP) bound to Cdc42. The essential residues are depicted in yellow, hydrophobic residues depicted in

purple and the red sites show positions where the new dyes were attached to generate changes in fluorescence when the CBD bound Cdc42.

FIG. 15 provides the intensity of fluorescence at various wavelengths observed when fluorescently labeled CBD binds to cdc42. When cdc42 is activated with GTP γ S, highly intense fluorescence is observed (line labeled "GTP γ S"), compared to when no cdc42 is present (line labeled "no cdc42"), or when cdc42 is not activated (line labeled "GDP").

FIG. 16 shows how the present methods can be used for in vitro assays on crude cellular lysates. In this case, the fluorescently labeled CBD was added to a neutrophil lysate. Upon binding to activated cdc42, CBD will emit fluorescence of greater intensity. At time zero fMLP, which stimulates neutrophils to activate cdc42, was added to the cellular lysate and the amount of fluorescence generated by the lysate (●) was measured as a function of time. As a control, the maximal amount of fluorescence that could be obtained from the lysate was estimated by adding saturating levels of GTP γ S (▼), which would activate most or all of the cdc42 in the cellular lysate. In this manner the levels of cdc42 in a cellular population can be quantified.

FIG. 17 provides photomicrographs of live cells injected with fluorescently labeled CBD. The lighter colors indicate where the activated cdc42 is present within the cells. The merocyanin dye ("mero") is an environmentally sensitive dye that will fluoresce at higher intensity upon binding of the CBD-mero conjugate to activated cdc42. In this case, the intensity of fluorescence emitted by a non-environmentally sensitive fluorophore (Alexa) linked to CBD was determined relative to the intensity of fluorescence emitted by the CBD-mero fluorophore. The ratio of the two permitted background fluorescence intensity to be subtracted so that the fluorescence from activated cdc-42 could be localized with greater precision.

FIG. 18 illustrates the affinity of the labeled leucine zipper peptide biosensor for the leucine peptide tag as determined by equilibrium fluorescence titration, which monitored the changes in anisotropy (○) and fluorescence intensity (●) as the amount of tag binding peptide was increased. Upon binding, the labeled biosensor showed a drop in quantum yield and a more than 230%

increase in rhodamine anisotropy, indicating considerable reduction in rotation of the dye. The anisotropy values plotted against the unlabeled peptide tag concentration were fit to an equation describing this interaction as a function of unlabeled peptide concentration, indicating a tight interaction with a dissociation constant (K_d) of 5.4 ± 1.1 nM.

FIG. 19 provides photomicrographs of the same cell with visualization of α -actinin by fluorescently labeled GFP- α -actinin and a rhodamine peptide biosensor of the invention. Cos-7 cells were transfected with an α -actinin construct with GFP fused to the N-terminus and the tag peptide on the C-terminus. The tag peptide and the rhodamine-peptide biosensor bound as illustrated in Figure 18. The rhodamine-peptide biosensor was injected into these cells and α -actinin localization was visualized using GFP or rhodamine fluorescence. Figures 19a and 19c show GFP images and Figures 19b and 19d show rhodamine images taken of the same cells. The similar fluorescence distribution in the GFP image (Figure 19a) and rhodamine image (Figure 19b) demonstrate labeling of α -actinin with a high degree of specificity. Figures 19c (GFP) and 19d (rhodamine) show deconvolution imaging of the same cell. This technique removes out-of-focus light, to demonstrate more clearly the coincidence of emission from the two fluorophores. (Bar = 10 microns).

FIG. 20 provides photomicrographs of a different cell than is shown in Figure 19, with visualization of α -actinin by fluorescently labeled GFP- α -actinin and a rhodamine peptide biosensor of the invention as described in Figure 19. Figure 20a provides GFP fluorescence and Figure 20b provides rhodamine fluorescence. (Bar = 10 microns).

FIG. 21 is a representative example from control experiments in which cells were transfected with the GFP construct described in Figure 18, but not injected with rhodamine peptide. Figure 21a shows GFP fluorescence and Figure 21b shows rhodamine fluorescence, taken under exposure conditions and with fluorescence levels similar to those in Figures 19 and 20. This control demonstrated that the rhodamine fluorescence was not simply 'bleedthrough' of GFP into the rhodamine image. Similar results were obtained by injecting rhodamine peptide biosensor into nontransfected cells, showing that the GFP

fluorescence was not due to rhodamine 'bleedthrough' (data not shown). (Bar = 10 microns).

FIG. 22 provides photomicrographs illustrating fluorescent labeling of an endoplasmic reticulum membrane protein *in vivo*. A subunit of the Fc receptor (Fc ϵ RI α), a protein which spans the endoplasmic reticulum membrane, was tagged with . Cells were cotransfected with an MHC-GFP fusion protein as an endoplasmic reticulum marker, and Fc receptor fused to the peptide tag. The cells were then injected with the rhodamine-labeled peptide and imaged in both rhodamine and GFP channels. Figure 22a shows the GFP fluorescence of the MHC marker, Figure 22c shows the rhodamine fluorescence of the tagged Fc ϵ RI α receptor. The images from Figures 22a and c are merged in Figure 22b. The MHC-GFP is more broadly distributed relative to the Fc receptor, which appears to concentrate in ER regions closer to the nucleus and in the nuclear membrane. Note that the GFP-expressing cell which was not injected with rhodamine does not artefactually appear in the rhodamine image, indicating again that similar rhodamine and GFP fluorescence are not due to 'bleedthrough' of GFP fluorescence into the rhodamine image (Bar = 10 microns).

FIG. 23 shows the fluorescence of the ERK2-dye construct under different conditions. ERK2 was activated with MEK for 0-60 min, as indicated in the figure. When Mg and ATP were added, ERK2 was phosphorylated; no such phosphorylation was observed when no Mg, ATP or MEK was present. The data shown here demonstrate that MEK can interact with the labeled protein.

FIG. 24 shows the fluorescence intensity as a function of wavelength for ERK2 when ERK2 is phosphorylated and unphosphorylated. Approximately 1.6-fold increase in emission was observed, when the Erk2 was phosphorylated.

FIG. 25 shows the fluorescence of ERK2, in the presence and absence of MEK, as a function of time of incubation with MEK. No substantial Erk2 fluorescence was observed when MEK was not present. This result demonstrated the suitability of the present methods for detecting activation of ERK2 in live cells.

Detailed Description

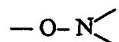
The following definitions are used, unless otherwise described.

Alkylene, alkenylene, alkynylene, etc. denote both straight and branched groups; but reference to an individual radical such as "propylene" embraces only
5 the straight chain radical; a branched chain isomer such as "isopropylene" being specifically referred to. Aryl denotes a phenyl radical or an ortho-fused bicyclic carbocyclic radical having about nine to ten ring atoms in which at least one ring is aromatic.

The term "amino acid," includes the residues of the natural amino acids
10 (e.g. Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Hyl, Hyp, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, and Val) in D or L form, as well as unnatural amino acids (e.g. phosphoserine, phosphothreonine, phosphotyrosine, hydroxyproline, gamma-carboxyglutamate; hippuric acid, octahydroindole-2-carboxylic acid, statine, 1,2,3,4,-tetrahydroisoquinoline-3-carboxylic acid, penicillamine,
15 ornithine, citruline, α -methylalanine, para-benzoylphenylalanine, phenylglycine, propargylglycine, sarcosine, and tert-butylglycine). The term also includes natural and unnatural amino acids bearing a conventional amino protecting group (e.g. acetyl or benzyloxycarbonyl), as well as natural and unnatural amino acids protected at the carboxy terminus (e.g. as a (C₁-C₆)alkyl, phenyl or benzyl ester
20 or amide). Other suitable amino and carboxy protecting groups are known to those skilled in the art (See for example, T.W. Greene, *Protecting Groups In Organic Synthesis*; Wiley: New York, 1981, and references cited therein).

The term "peptide" includes any sequence of 2 or more amino acids. The sequence may be linear or cyclic. For example, a cyclic peptide can be prepared
25 or may result from the formation of disulfide bridges between two cysteine residues in a sequence. Thus, the term includes proteins, enzymes, antibodies, oligopeptides, and polypeptides. Peptide sequences specifically recited herein are written with the amino terminus on the left and the carboxy terminus on the right.

30 An "aminoxy group" is a group having the following formula



wherein the open valences are filled by any acceptable radical. A "secondary aminoxy group" is an aminoxy group where one of the open valences on the nitrogen is filled by a radical other than a hydrogen.

5 The term "functional molecule" includes any compound that can be linked to a peptide to provide a peptide conjugate having useful properties. Such conjugates may be useful for studying the structure or function of the peptide, or a polypeptide, antibody, antigen or other protein. Such conjugates can also be used for therapeutic treatments and diagnoses. Functional molecules linked to
10 peptides by the present methods can be used in any assay, procedure or tracing protocol known to one of skill in the art. The functional molecules may also be used with biosensors as contemplated below. Peptide-conjugates with such functional molecules may be useful for drug screening, as pharmacological tools, as research tools, or as therapeutic agents. For example, the term functional
15 molecule includes labels, dyes, ESR probes, reporting groups, biophysical probes, peptides, polynucleotides, therapeutic agents, pharmaceuticals, toxins, cross-linking groups (chemical or photochemical), a compound that modifies the biological activity of the peptide, or a caged molecule (e.g. a reporting molecule or a biologically active agent that is masked and that can be unmasked by
20 photoactivation or chemical means).

The term "biophysical probe" includes any group that can be detected *in vitro* or *in vivo*, such as, for example, a fluorescent group, a phosphorescent group, a nucleic acid indicator, an ESR probe, another reporting group, a moiety, or a dye that is sensitive to pH change, ligand binding, or other environmental
25 aspects.

Amino acids and peptides that include one or more aminoxy groups are useful intermediates for preparing peptide conjugates. The aminoxy group(s) can typically be positioned at any suitable position on the amino acid or peptide. For example, the aminoxy group(s) can conveniently be incorporated into the
30 side chain of the amino acid or into one or more side chains of the peptide. Thus, as used herein with respect to the amino acids and peptides of the invention, the term "a radical comprising one or more aminoxy groups"

includes any organic group that can be attached to the amino acid or peptide that includes one or more aminoxy groups. For example, the term includes a carbon chain having two to ten carbon atoms; which is optionally partially unsaturated (i.e. contains one or more double or triple bonds); which chain is optionally 5 interrupted by one or more (e.g. 1, 2, or 3) -NH-, -O-, or -S-; which chain is optionally substituted on carbon with one or more (e.g. 1, 2, or 3) oxo (=O) groups; and which chain is optionally substituted with one or more (e.g. 1, 2, or 3) aminoxy groups. Preferably, the aminoxy group(s) are secondary aminoxy groups.

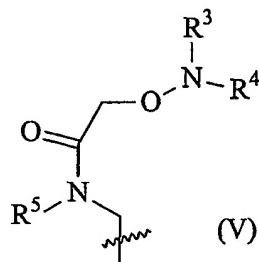
10 The term "cross-linking group" refers to any functionality that can form a bond with another functionality, such as photoaffinity label or a chemical crosslinking agent.

15 The term "caged molecule" includes a molecule or reporter group that is masked such that it can be activated (i.e. unmasked) at a given time or location of choice, for example using light or a chemical agent.

Specific and preferred values listed below for radicals, substituents, and ranges, are for illustration only; they do not exclude other defined values or other values within defined ranges for the radicals and substituents.

20 Specifically, (C₁-C₆)alkylene can be methylene, ethylene, propylene, isopropylene, butylene, iso-butylene, sec-butylene, pentylene, or hexylene; (C₃-C₈)cycloalkyl can be cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, or cyclooctyl; (C₂-C₆)alkenylene can be vinylene, allylene, 1-propenylene, 2-propenylene, 1-butenylene, 2-butenylene, 3-butenylene, 1-pentenylene, 2-pentenylene, 3-pentenylene, 4-pentenylene, 1-hexenylene, 2-hexenylene, 3-hexenylene, 4-hexenylene, or 5-hexenylene; (C₂-C₆)alkynylene can be ethynylene, 1-propynylene, 2-propynylene, 1-butynylene, 2-butynylene, 3-butynylene, 1-pentynylene, 2-pentynylene, 3-pentynylene, 4-pentynylene, 1-hexynylene, 2-hexynylene, 3-hexynylene, 4-hexynylene, or 5-hexynylene; and aryl can be phenyl, indenyl, or naphthyl;

30 A specific value for R is a radical of formula (V):



wherein

R^3 is hydrogen, (C_1-C_6)alkyl, an amino protecting group, or a radical comprising one or more aminoxy groups;

5 R^4 is hydrogen, or an amino protecting group; and

R^5 is hydrogen, or (C_1-C_6)alkyl.

A specific value for R^1 is hydrogen or benzyloxycarbonyl.

A specific value for R^2 is hydrogen.

A specific value for R^3 is methyl.

10 A specific value for R^4 is hydrogen, 2-chlorobenzyloxycarbonyl, or benzyloxycarbonyl.

A specific value for R^5 is hydrogen.

A specific value for R^6 is an antibody.

15 A specific value for R^6 is a peptide or polypeptide or antibody that includes about 2 to about 1000 amino acids. A more specific value for R^6 is a peptide that includes about 5 to about 500 amino acids. An even more specific value for R^6 is a peptide that includes about 10 to about 100 amino acids.

Specifically X is a linking group that is about 5 angstroms to about 100 angstroms in length. More specifically, X is a linking group of about 5 angstroms to about 25 angstroms in length.

20 Specifically X is $-R_a-C(=O)-NH-R_b-$ wherein each of R_a and R_b is independently (C_1-C_6)alkylene. Preferably, each of R_a and R_b is methylene ($-CH_2-$).

A preferred value for R^6 is KKKEKERPEISLPSDFEHTIHVG

25 DACTGEFTGMPEQWARLLQT (SEQ ID NO: 1) or an antibody.

A specific value for R^7 is hydrogen.

Another specific value for R⁷ is (C₁-C₆)alkyl.

A preferred value for R⁷ is methyl.

Specifically Y is a linking group that is about 5 angstroms to about 100 angstroms in length. More specifically, Y is a linking group of about 5
5 angstroms to about 25 angstroms in length.

A specific value for Y is (C₁-C₆)alkylene.

A preferred value for Y is methylene (-CH₂-).

Fluorescent Dyes

10 Any fluorescent dye known to one of skill in the art is contemplated by the present invention as a functional molecule. A fluorescent dye can be excited to fluoresce by exposure to a certain wavelength of light. When used as the reporter molecule of the biosensors of the present invention, the dye is preferably environmentally sensitive. As used herein, "environmentally sensitive" means
15 that the signal from the functional molecule changes when the peptide, polypeptide or antibody interacts with, or becomes exposed to, a different environment. For example, when the environmentally sensitive functional molecule is a fluorescent dye, the fluorescence from that fluorescent dye will change as the environment changes. In one embodiment an environmentally
20 sensitive fluorescent dye attached to a peptide, polypeptide or antibody will fluoresce differently upon target binding by the peptide, polypeptide or antibody to which the dye is attached. Any dye which emits fluorescence and whose fluorescence changes when the pH or the hydrophilicity/hydrophobicity of the environment changes is an environmentally sensitive dye contemplated by the
25 present invention.

Preferred fluorescent groups include molecules that are capable of absorbing radiation at one wavelength and emitting radiation at a longer wavelength, such as, for example, Alexa-532, Hydroxycoumarin, Aminocoumarin, Methoxycoumarin, Coumarin, Cascade Blue, Lucifer Yellow,
30 P-Phycoerythrin, R-Phycoerythrin, (PE), PE-Cy5 conjugates, PE-Cy7 conjugates, Red 613, Fluorescein, BODIPY-FL, BODIPY TR, BODIPY TMR, Cy3, TRITC, X-Rhodamine, Lissamine Rhodamine B, PerCP, Texas Red, Cy5,

- Cy7, Allophycocyanin (APC), TruRed, APC-Cy7 conjugates, Oregon Green, Tetramethylrhodamine, Dansyl, Dansyl aziridine, Indo-1, Fura-2, FM 1-43, DilC18(3), Carboxy-SNARF-1, NBD, Indo-1, Fluo-3, DCFH, DHR, SNARF, Monochlorobimane, Calcein, N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl) amine
5 (NBD), ananilinonaphthalene, deproxyl, phthalamide, amino pH phthalamide, dimethylamino-naphthalenesulfonamide, probes comparable to Prodan, Lordan or Acrylodan and derivatives thereof. Coumarin fluorescent dyes include, for example, amino methylcoumarin, 7-diethylamino-3-(4'-(1-maleimidyl)phenyl)-4-methylcoumarin (CPM) and N-(2-(1-maleimidyl)ethyl)7-
10 diethylaminocoumarin-3-carboxamide (MDCC). Preferred fluorescent probes are sensitive to the polarity of the local environment and are available to those of skill in the art.

Other useful functional molecules include those that display fluorescence resonance energy transfer (FRET). Many such donor-acceptor pairs 15 are known, and include fluorescein to rhodamine, coumarin to fluorescein or rhodamine, etc. Still another class of useful label pairs include fluorophore-quencher pairs in which the second group is a quencher, which decreases the fluorescence intensity of the fluorescent group. Some known quenchers include acrylamide groups, heavy atoms such as iodide and bromate, nitroxide spin 20 labels such as TEMPO, etc. These can be adapted for use as environmentally sensitive functional molecules of biosensors.

Exemplary fluorescent proteins which can be used to label the present peptides, polypeptides and antibodies include green fluorescent protein (GFP), cyan fluorescent protein (CFP), red fluorescent protein (RFP), yellow fluorescent 25 protein (YFP), enhanced GFP (EGFP), enhanced YFP (EYFP), and the like.

New Fluorescent Dyes

The present invention also provides novel fluorescent dyes that retain high fluorescence emission after conjugation to other molecules and avoid 30 problems of aggregation and insolubility. These dyes are particularly preferred for many of the imaging methods and conjugates contemplated but need not be

restricted to use in the methods and conjugates contemplated herein. Thus, the present invention is directed to highly fluorescent dyes of structure IV wherein,

each m is separately an integer ranging from 1-3;

n is an integer ranging from 0 to 5;

5 R⁸, R¹¹ and R¹² are separately CO, SO₂, C=C(CN)₂, S, O or C(CH₃)₂;

each R¹³ is alkyl, branched alkyl or heterocyclic ring derivatized with charged groups to enhance water solubility and enhance photostability;

10 R⁹ and R¹⁰ are chains carrying charged groups to enhance water solubility (i.e. sulfonate, amide, ether) and/or chains bearing reactive groups for conjugation to other molecules. The reactive group is a functional group that is chemically reactive (or that can be made chemically reactive) with functional groups typically found in biological materials, or functional groups that can be readily converted to chemically reactive derivatives using methods well known in the art. In one embodiment of the invention, the charged and reactive groups are separately haloacetamide (-NH-(C=O)-CH₂-X), where X is Cl, Br or I.

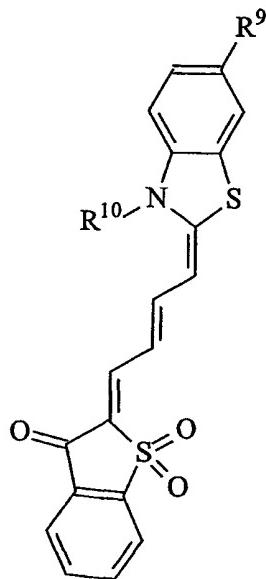
15 Alternatively, the charged and reactive groups are separately amine, maleimide, isocyanato (-N=C=O), isothiocyanato(-N=C=S), acyl halide, succinimidyl ester, or sulfosuccinimidyl ester. In another embodiment, the charged and reactive groups are carboxylic acid (COOH), or derivatives of a carboxylic acid. An appropriate derivative of a carboxylic acid includes an alkali or alkaline earth metal salt of carboxylic acid. Alternatively, the charged and reactive groups are reactive derivatives of a carboxylic acid (-COORx), where the reactive group Rx is one that activates the carbonyl group of -COORx toward nucleophilic displacement. In particular, Rx is any group that activates the carbonyl towards nucleophilic displacement without being incorporated into the final displacement product. Examples of COORx : ester of phenol or naphtol that is further substituted by at least one strong electron withdrawing group, or carboxylic acid activated by carbodiimide, or acyl chloride, or succinimidyl or sulfosuccinimidyl ester. Additional charged and reactive groups include, among others, sulfonyl halides, sulfonyl azides, alcohols, thiols, semicarbazides, hydrazines or hydroxylamines.

For all dyes of the invention, any net positive or negative charges possessed by the dye are balanced by a biologically compatible counterion or counterions. As used herein, a substance that is biologically compatible is not toxic as used, and does not have a substantially deleterious effect on

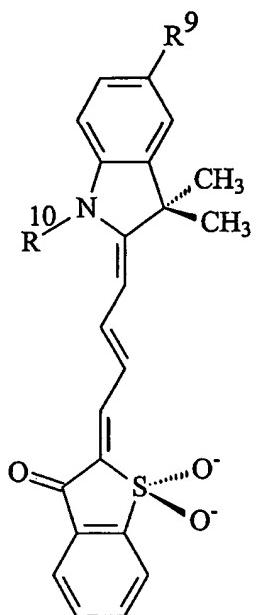
- 5 biomolecules. Examples of useful counterions for dyes having a net negative charge include, but are not limited to, alkali metal ions alkaline earth metal ions, transition metal ions, ammonium and substituted ammonium ions. Examples of useful counterions for dyes having a net positive charge include, but are not limited to chloride, bromide, iodide, sulfate, phosphate, perchlorate, nitrate,
- 10 tetrafluoroborate.

As used herein, R⁹ and R¹⁰ chains for conjugation are alkyl, of unlimited length, preferably with 1-6 carbons, and can include other moieties such as ether, amide, or sulfonate to improve water solubility. The groups can be substituted on the end or in the middle of the alkyl chain.

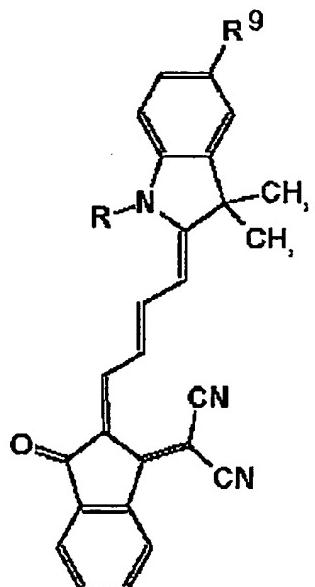
- 15 In one embodiment, the fluorescent dye of the present invention has the following structure (VI):



- 20 In another embodiment, the fluorescent dye of the present invention has the following structure (VII):



In another embodiment, the fluorescent dye of the present invention has the following structure (VIII):

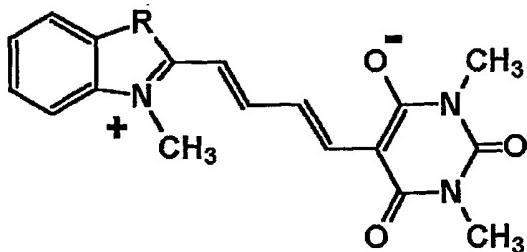


5

Depending upon their environment, the fluorescent dyes of the present invention can exist in somewhat different polarization states. This property can

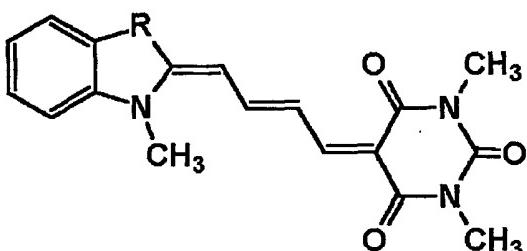
modulate the solubility and the emission wavelength of the dye. For example, the fluorescent dye depicted below can be charged or non-charged.

5



10

15



The fluorescent dyes of the present invention can absorb and emit light at a variety of wavelengths, depending on the arrangement and variety of substituents employed. Thus, the choice of whether to use R⁸, R¹¹ and R¹² as -CO-, -SO₂-, 20 -C=C(CN)₂- , -S-, -O- or -C(CH₃)₂- can influence the absorption and emission wavelength. However, by varying the substituents of the present invention, one of skill in the art can readily ascertain which combination of substituents will yield a fluorescent dye with a desired absorption and emission spectrum.

Moreover, according to the present invention, the degree of conjugation of the dye, and in particular, the length of the alkylene chain connecting the two rings, can predictably influence the absorption and emission wavelength of the dye. Thus, addition of one -C=C- group can shift the fluorescence wavelength about +100 nm. Smaller incremental changes in the emission wavelength can be 30 made by adding a conjugated group to one of the rings in the dye. Thus, one of skill in the art can readily modulate the emission wavelength of the dye as desired. In one embodiment, the absorption and emission wavelengths can be

altered to range from about 300 nm to about 800 nm. Preferably, the absorption and emission wavelengths of the present dyes range from about 450 nm to higher wavelengths. Any variety of methods can be used to make the present dyes.

- Preferred nucleic acid indicators include intercalating agents and
- 5 oligonucleotide strands, such as, for example, YOYO-1, Propidium Iodide, Hoechst 33342, DAPI, Hoercht 33258, SYTOX Blue, Chromomycin A3, Mithramycin, SYTOX Green, SYTX Orange, Ethidium Bromide, 7-AAD, Acridine Orange, TOTO-1, TO-PRO-1, Thiazole Orange, Propidium Iodide, TOTO-3, TO-PRO-3, LDS 751.

10

Synthon and Peptide Intermediates

The synthetic intermediates (i.e. synthons) of the invention that include one or more aminoxy groups can be incorporated into peptides using a variety of techniques that are known in the art. For example, as discussed below, the

15 synthons can be incorporated into a peptide using solid-phase peptide synthesis, solution-phase peptide synthesis, native chemical ligation, intein-mediated protein ligation, and chemical ligation.

Peptides may be prepared using solid-phase peptide synthesis (SPPS). For example, according to the SPPS technique, protected amino acids in organic

20 solvents can be added one at a time to a resin-bound peptide chain, resulting in the assembly of a target peptide having a specific sequence in fully-protected, resin-bound form. The product peptide can then be released by deprotection and cleavage from the resin support (Wade, L.G., JR., Organic Chemistry 4th Ed. (1999)). As illustrated in Example 2 below, amino acids containing an aminoxy

25 functional group can be incorporated into peptides using SPPS. Use of this methodology allows an amino acid containing an aminoxy functional group to be positioned at a desired location within a synthesized peptide chain.

Amino acids containing an aminoxy group can also be incorporated into a peptide using solution-phase peptide synthesis (Wade, L.G., JR. Organic

30 Chemistry 4th Ed. (1999)). Solution-phase peptide synthesis involves protecting the amino-terminus of a peptide chain followed by activation of the carboxyl-

terminus allowing the addition of an amino acid or a peptide chain to the carboxy-terminus (Wade, L.G., JR. Organic Chemistry 4th Ed. (1999)).

Native chemical ligation is a procedure that can be used to join two peptides or polypeptides together thereby producing a single peptide or 5 polypeptide having a native backbone structure. Native chemical ligation is typically carried out by mixing a first peptide with a carboxy-terminal α -thioester and a second polypeptide with an amino-terminal cysteine (Dawson, P.E., et al., (1994), Science 266:776-779; Cotton, G.J., et al., (1999), J. Am. Chem. Soc. 121:1100-1101). The thioester of the first peptide undergoes 10 nucleophilic attack by the side chain of the cysteine residue at the amino terminus of the second peptide. The initial thioester ligation product then undergoes a rapid intramolecular reaction because of the favorable geometric arrangement of the alpha-amino group of the second peptide. This yields a product with a native peptide bond at the ligation site. A polypeptide beginning 15 with cysteine can be chemically synthesized or generated by intein vectors, proteolysis, or cellular processing of the initiating methionine. This method allows mixing and matching of chemically synthesized polypeptide segments.

The synthons of the invention are particularly useful in combination with native chemical ligation, because native chemical ligation allows a synthetic 20 peptide having a specifically positioned amino acid (e.g. a synthon of the invention) to be selectively ligated to other peptides or into a larger polypeptide, antibody or protein. The ability to specifically incorporate aminoxy modified amino acids into a peptide chain allows useful moieties to be linked at any 25 position within a peptide, polypeptide, antibody or protein. Examples of such moieties that can be incorporated into a peptide using this method include, but are not limited to, phosphorylated or glycosylated amino acids, unnatural amino acids, tags, labels, crosslinking reagents, biosensors, reactive groups, and fluorophores. Another advantage of native chemical ligation is that it allows incorporation of peptides into a polypeptide that are unable to be added by 30 ribosomal biosynthesis.

Intein-mediated protein ligation may also be used to selectively place amino acids containing aminoxy functional groups into peptides. Inteins are

intervening sequences that are excised from precursor proteins by a self-catalytic mechanism and thereby expose reactive ends of a peptide. Intein vectors have been developed that not only allow single-step purification of proteins, but also yield polypeptides with reactive ends necessary for intein-mediated protein

- 5 ligation (IPL) (also called expressed protein ligation)(EPL) (Perler, F.R. and Adam, E., (2000) Curr. Opin. Biotechnol. 11(4):377-83; and Evans, T.C., et al., (1998) Protein Sci 7:2256-2264). This method allows a peptide having a selectively placed amino acid containing an aminoxy functional group to be readily ligated to any peptide with reactive ends generated by intein excision.

- 10 Two peptides or polypeptides may also be linked through use of chemical ligation. Chemical ligation occurs when two peptide segments are each linked to functional groups that react with each other to form a covalent bond producing a non-peptide bond at the ligation site (Wilken, J. and Kent, S.B.H., (1998) Curr. Opin. Biotechnol. 9:412-426). This method can be used to ligate a peptide
- 15 having a specifically positioned aminoxy functional group to another peptide or polypeptide to produce a desired polypeptide that may be later linked to a detectable group.

- A functional molecule ("D") can be attached to a peptide comprising an aminoxy group through a direct linkage (e.g. an amide bond -O-N-C(=O)-D) or
- 20 through a linking group. The structure of the linking group is not crucial, provided it does not interfere with the use of the resulting labeled peptide. Preferred linking groups include linkers that separate the aminoxy nitrogen and the detectable group by about 5 angstroms to about 100 angstroms. Other preferred linking groups separate the aminoxy nitrogen and the detectable
- 25 group by about 5 angstroms to about 25 angstroms.

- For example, the linking group can conveniently be linked to the detectable group through an: 1) amide (-N(H)C(=O)-, -C(=O)N(H)-), 2) ester (-OC(=O)-, -C(=O)O-), 3) ether (-O-), 4) thioether (-S-), 5) sulfinyl (-S(O)-), or 6) sulfonyl (-S(O)₂) linkage. Such a linkage can be formed from suitably
- 30 functionalised starting materials using synthetic procedures that are known in the art.

The linking group can conveniently be linked to the nitrogen of the aminoxy group to form an amide (-O-N(H)C(=O)- or a thiourea linkage (-O-N-C(=S)-N-) linkage, using reagents and conditions that are known in the art.

The aminoxy group can be attached to a peptide through a direct bond

- 5 (e.g. a carbon-oxygen bond) between the aminoxy oxygen and a side chain of the peptide, or the aminoxy group can be attached to the peptide through a linking group. The structure of the linking group is not crucial, provided it does not interfere with the use of the resulting labeled peptide. Preferred linking groups include linking groups that separate the aminoxy oxygen and the side
- 10 chain of the peptide by about 5 angstroms to about 100 angstroms. Other preferred linking groups separate the aminoxy oxygen and the side chain of the peptide by about 5 angstroms to about 25 angstroms.

A specific linking group (e.g. X or Y) can be a divalent (C_1 - C_6)alkylene, (C_2 - C_6)alkenylene, or (C_2 - C_6)alkynylene chain, or a divalent (C_3 - C_8)cycloalkyl, 15 or aryl ring.

Thus, a simple and efficient synthetic methodology for site-specific labeling of peptides after synthesis has been developed that provides high yield, selectivity and compatibility with both solid-phase peptide synthesis and C^α -thioester peptides. The approach and primary advantages can be summarized 20 as follows:

- (1) A protected aminoxy amino acid that can be incorporated into peptides has been synthesized;
- (2) Procedures have been optimized to yield highly efficient and specific modification of the aminoxy nitrogen in the presence of unprotected competing 25 nucleophiles, including cysteine, lysine and amino groups;
- (3) One preferred electrophile that can be used for labeling, an activated carboxylic ester, is readily available in the majority of commercially available fluorescent dyes and labels;
- (4) Labeling of the aminoxy group occurs after synthesis and 30 purification, thus enabling the use of chemically sensitive fluorophores and labels that would otherwise not survive earlier synthetic procedures;

(5) The synthetic methodology is compatible with the steps required for the synthesis of proteins by total chemical synthesis or expressed protein ligation, namely synthesis of C^α-peptide thioesters and amide-forming ligations; and

5 (6) Combinatorial screening of both the functional molecule and its placement will enable the rapid synthesis of optimally labeled polypeptide-based biosensors.

Biosensors

10 Using the methods of the present invention with the synthons and peptides provided herein, antibodies, antigens and other polypeptides can be labeled with or attached to functional molecules. Such labeled or conjugated antibodies and polypeptides have particular utility as biosensors. As used herein, a "biosensor" is a peptide, antigen, polypeptide or antibody with an attached 15 functional molecule. In one embodiment, the functional molecule is a label or dye, however, as used herein, a biosensor can have any functional molecule attached thereto. For example, the functional molecule can be a label, dye, biophysical probe, peptide, polynucleotide, therapeutic agent, pharmaceutical, toxin, cross-linking group (chemical or photochemical), a compound that 20 modifies the biological activity of the peptide, or a caged molecule (e.g. a reporting molecule or a biologically active agent that is masked and that can be unmasked by photoactivation or chemical means).

In one embodiment the functional molecule is dye or label. Such a dye or label can be an environmentally-sensitive fluorescent dye such that 25 fluorescence is emitted by the dye can change when the protein biosensor becomes exposed to a different environment. Such a change in environment can occur, for example, when the biosensor binds to, or associates with, a target protein or cellular structure. The change in fluorescence can be used to quantify or otherwise monitor the amount of binding or interaction between the biosensor 30 and the target site. The Examples provided herein further illustrate how such a biosensor can be used.

The present invention provides methods for identifying an optimal position for the functional molecule on the peptide, polypeptide or antibody which involve generating a series of peptides, each peptide have the same amino acid sequence and functional molecule. However, the functional molecule is

5 positioned at a different location along the backbone of each peptide in the series. To determine which location is best, the strength of the signal from the various peptides is observed under different conditions and the optimal location provides optimal functioning by the functional molecule and the peptide, polypeptide or antibody. For example, when the functional molecule provides a

10 signal, a stronger signal is preferred so long as the function and the chemical and physical properties of the peptide, protein or antibody are not impaired. When an environmentally sensitive functional molecule is chosen, a maximal change in signal is preferred as the environment is changed. For example, when the environmentally sensitive functional molecule is attached to detect an interaction

15 of the peptide with a target, a maximal change in signal is preferred when the peptide binds or interacts with its target, unless of the location of the functional molecule affects the binding affinity, binding selectivity or another desirable attribute of the peptide.

Thus to determine an optimal location for a functional molecule in an

20 antibody or polypeptide which can bind to a target, a series of peptides are first synthesized, each with the functional molecule at a different position. The peptides can be incorporated into the polypeptide or antibody using the methods described herein to generate a series of biosensors, each with a functional molecule at a somewhat different position. The interaction of the different

25 polypeptide or antibody biosensors with target is observed. In general, an optimal position for the functional molecule on such a biosensor is that position which permits stable and selective binding to target with a maximal signal change upon binding.

However, some variability in binding and signal strength can be tolerated

30 so long as an observable, localized signal is detected upon interaction of the biosensor and the target. Thus, the strength of binding between the biosensor and target should be sufficient to permit observation of bound biosensor. If the

biosensor is only transiently bound, little or no localized signal from the probe may be observed; instead, only diffuse signal from biosensor in solution may be observed. Similarly, if the biosensor is bound non-selectively, non-localized signal may be detected from many sites. Obviously, a strongly localized signal
5 that clearly correlates with biosensor binding to target is preferred. A readily detectable change in signal strength or quality as the biosensor and target interact is also preferred.

Procedures known to one of skill in the art can be used to detect a signal provided by the functional molecule and to correlate a change in signal by an
10 environmentally sensitive functional molecule. Signals contemplated by the present invention include fluorescent emissions, radioactive emissions, enzymatic production of a colored product, and the like. One of skill in the art can readily detect these signals using a fluorescence microscope, a scintillation counter, a light or radioactively sensitive photoemulsion, a light microscope, a
15 spectrophotometer or other means. When an environmentally-sensitive functional molecule is used, the signal can change in lifetime, strength, color or other quality to signal interaction of the functional molecule with the environment. For example, when the environmentally-sensitive functional molecule is a fluorescent dye, the signal change can be a color, wavelength,
20 intensity or lifetime of fluorescence emitted by the dye. The change need only be detectable, for example, a change in wavelength of fluorescence of about 50 nm to about 300 nm can be readily detected. However, a change in wavelength of greater than about 10 nm is preferred. More preferably the change in wavelength is greater than 20 nm. In one embodiment the change in wavelength
25 can vary from about 30 nm to about 200 nm.

Where convenient, a peptide, as opposed to a polypeptide or antibody, with an attached functional molecule may be a biosensor, for example, when the binding affinity and selectivity of the peptide is representative of the larger protein of which it is normally a part. Alternatively, the labeled peptide can be
30 ligated into a polypeptide to form the protein or antibody of which it is normally a part using the methods described herein. In one embodiment, the peptide is incorporated onto one of the termini of a protein, for example, through

procedures described in Curr. Op. Biotechnology 9:412 (1998) or Ann. Rev. Biochem. 57:957-89 (1988). In another embodiment, the peptide can be incorporated into the middle of a protein, for example, by using the procedures described in PNAS 95:6705 (1998).

5

Targets

- As used herein, a "target" is any molecule, structure or complex that a peptide, polypeptide or antibody linked to a functional molecule by the present methods can interact with. Targets contemplated by the present invention
- 10 include antigens, antibodies, proteins, enzymes, membrane proteins, endoplasmic reticulum proteins, structural proteins, major histocompatibility proteins, DNA binding proteins, receptors, ligands, cofactors, nucleic acids, kinases, GTPases, ATPases, proteins involved in motility and the like. Targets may undergo structural changes and other types of changes, for example,
 - 15 phosphorylation, de-phosphorylation, conformational changes, ligand binding, co-factor binding, activation, post-translational modification, carbohydrate or sugar attachment, membrane interactions and the like. Specific targets include calmodulin, Rho GTPases, rac, cdc42, mitogen-activated protein kinase (MAP kinase), Erk1, Erk2, Erk3, Erk4, IgE receptor (F_c RI) actin, α -actinin, myosin,
 - 20 and major histocompatibility proteins.

Targets can have "tag" sequences that permit binding of a biosensor or labeled polypeptide to the target. As used herein, a "tag" is any binding site for a biosensor. In general, a tag is a peptide or polypeptide segment that is recognized by the biosensor and to which the biosensor will bind with sufficient binding affinity to permit detection and/or visualization of the target. Any protein binding domain or ligand binding site of a receptor can be used. Thus, a tag can be, for example, an antigenic epitope to which an antibody can bind, a part of a leucine-zipper to which another part of a leucine zipper can bind, a domain of a homeotic protein that mediates binding to a DNA site, receptor-ligand binding site.

The dimerizing domain from yeast GCN4 protein can also be used to form a tag-biosensor dimer. The GCN4 dimerizing domain is a single α -helix

which pairs with its partner to form a parallel coiled-coil (O'Shea, Rutkowski, Stafford and Kim, *Science*, 245, 646, (1989)). The X-ray crystal structure of the GCN4 coiled-coil has been determined (O'Shea, Klemm, Kim and Alber, *Science*, 254, 539, (1991)), and factors affecting stability and oligomerization
5 state of native and mutant GCN4 coiled-coil peptides have been determined (Harbury, Zhang, Kim and Alber, *Science*, 262, 1401, (1993); O'Neil, Hoess and DeGrado, *Science*, 249, 774, (1990)). In addition, several model coiled-coil peptides have been extensively studied and rules governing stability of these designed peptides have been reported (Hodges, Zhou, Kay and Semchuk,
10 Peptide Res., 3, 123, (1990); Zhu et al., *Int. J. Peptide Protein Res.*, 40, 171, (1992); Lumb and Kim, *Biochemistry*, 34, 8642, (1995)).

Another useful tag-biosensor dimer can be formed from the binding domain from the Arc repressor of bacteriophage P22 (Schildbach, J F, Milla M E, Jeffrey PD, Raumann BE, Sauer RT (1995). *Biochemistry* 34: 13914-19). Arc repressor is a dimeric polypeptide of 53 amino acids. A tag-biosensor dimer can also be formed from the C-terminal tetramerizing domain from the tumor suppressor p53. The structure of this domain has been determined from the crystal Jeffrey, Gorina and Pavletich, *Science*, 267, 1498, (1995)) and in solution (Clore et al., *Nature Struct. Biol.*, 2, 321, (1995)). The Mnt repressor of
15 bacteriophage P22 (Waldburger, C. D. and R. T. Sauer (1995). *Biochemistry* 34(40): 13109-13116), the Mnt repressor polypeptide and streptavidin-avidin can be used to form tag-biosensor dimers. Streptavidin (Argarana, C., Kuntz, I.D., Birken, S, Axel, R. and Cantor, C. R. (1986) *Nucleic Acids Res.* 14:1871) and avidin (Green, N. M. (1975) *Adv. Protein Chem.* 29:85) are homologous
20 tetrameric polypeptides of approximately 125-127 and 128 amino acids, respectively. Biotin ligase (BLS) can also be used to make a tag-biosensor dimer. BLS polypeptide sequences have been described for various proteins: for example, the C-terminal 87 residues of the biotin carboxy carrier protein of
25 Escherichia coli acetyl-CoA carboxylase (Chapman-Smith, A., D. L. Turner, et al. (1994). *Biochem J.* 302: 881-7). The C-terminal 67 residues of carboxyl-terminal fragments of human propionyl-CoA carboxylase alpha subunit may be used (Leon-Del-Rio, A. and R. A. Gravel (1994) *J. Biol. Chem.* 269(37): 22964-

8); and residues 18-123 of Propionibacterium freudenreichii transcarboxylase 1.3S biotin subunit (Yamano, N., Y. Kawata, et al. (1992) Biosci. Biotechnol. Biochem. 56(7): 1017-1026).

- Such tag sequences can be present naturally or can be added to the target
- 5 by methods available to one of skill in the art. Different targets within a single cell can have different tags to which different biosensors can bind. This permits visualization of two or more targets within the same cell so that the interaction and dynamic interplay between targets can be observed in relation to other cellular structures. As described herein, biosensors can also be designed to
- 10 undergo fluorescence resonance energy transfer (FRET) when they become juxtaposed at an appropriate distance. Similarly, a target can be engineered to have two different tags so that two different biosensors can bind to the same target. Tag sequences can be added to a target protein, for example, by protein ligation procedures described herein or by standard molecular biology
- 15 techniques where a peptide or polypeptide is fused to another peptide or polypeptide. See, e.g., Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.).

20 **Antigen and Antibody Targets**

When antibodies are linked to functional molecules using methods of the present invention the target can be an antigen. Alternatively, an antigen or a peptide epitope can serve as a biosensor for an antibody target. The present invention contemplates use or detection of any antigen or antibody known to one

25 of skill in the art as a target.

In general, a molecule must have sufficient complexity and sufficient molecular weight in order to act as an antigen. In order to have sufficient complexity, the antigen must have at least one epitope. As used in this invention, the term "epitope" is meant to include any determinant capable of

30 specific interaction with the monoclonal antibodies of the invention. Epitopic determinants usually consist of chemically active surface groupings of molecules

such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics.

In order to have a sufficient molecular weight, the antigen generally must have a molecular weight that is greater than 2,000 daltons. Formerly, it was
5 thought that the lower molecular weight limit to confer antigenicity was about 5,000 daltons. However, antigenicity has recently been demonstrated with molecules having molecular weights as low as 2,000 daltons. Molecular weights of 3,000 daltons and more appear to be more realistic as a lower limit for immunogenicity, and approximately 6,000 daltons or more is preferred.

10 In preparing antigens to produce antibodies for attachment to functional molecule, it is desirable to use antigens with a high degree of purity. Accordingly, it is desirable to use a purification process permitting isolation of the antigen from antigenically distinct materials. Antigenically distinct materials are undesired large molecules that may compete with the target antigen for
15 antibody production thereby minimizing production of the desired antibodies or inducing cross-reactive antibodies of low specificity or affinity. The practice of the invention can accordingly include a number of purification steps using available techniques. Purification can, for example, be effected by size exclusion chromatography, ion exchange chromatography, dialysis, cold organic
20 solvent extraction, gel electrophoresis and/or fractional crystallization means which are available to one of skill in the art. Preferably an antigen used for antibody preparation is at least about 90% pure and more preferably at least about 99% pure.

25 Removal of small molecule reactants and reaction products from the synthesized antigen is generally desirable. However, some small molecular substances may be useful, for example, for control of pH and salinity. Thus, a convenient end-product form in which to recover the antigen is, in a buffered aqueous solution that is suitable for direct administration to animals.

30 Antibodies

The present invention contemplates linkage of any available antibody to functional molecules. Such antibodies can be polyclonal or monoclonal

antibodies. The term "antibody" as used in this invention is meant to include intact antibody molecules as well as fragments thereof, such as, for example, Fab and F(ab')₂, which are capable of binding to the antigen or its epitopic determinant.

5 Polyclonal antibodies can be raised by administration of an antigen of the invention to vertebrate animals, especially mammals such as goats, rabbits, rats or mice using known immunization procedures. Usually a buffered solution of the antigen accompanied by Freund's adjuvant is injected subcutaneously at multiple sites. A number of such administrations at intervals of days or weeks is
10 usually necessary. A number of animals, for example from 3 to 20, are injected with the expectation that only a small proportion will produce desirable antibodies. However, one animal can provide antibodies sufficient for thousands of assays. Antibodies are recovered from animals after some weeks or months.

Exemplary immunogenic carrier materials can be used with the antigen to
15 enhance the immune response. The carrier material can be a natural or synthetic substance, provided that it is an antigen or a partial antigen. For example, the carrier material can be a protein, a glycoprotein, a nucleoprotein, a polypeptide, a polysaccharide, a lipopolysaccharide, or a polyamino acid. *See also*, the carrier molecules and antibody production methods set forth in Cremer et al., "Methods
20 in Immunology" (1963), W. A. Benjamin Inc., New York, pp. 65-113 and Harlow and Lane, "Antibodies: A Laboratory Manual" (1988), Cold Spring Harbor Laboratory, New York, p.5. These disclosures are herein incorporated by reference.

A preferred class of natural carrier materials is the proteins. Proteins can
25 be expected to have a molecular weight in excess of 5,000 daltons, commonly in the range of from 34,000 to 5,000,000 daltons. Specific examples of such natural proteins are hen ovalbumin (OA), bovine serum albumin (BSA), keyhole limpet hemocyanin (KLH), horse gammaglobulin (HGG), and thyroglobulin.

Exemplary of synthetic carrier is the polyamino acid, polylysine. Where
30 the synthetic antigen comprises a partially antigenic carrier conjugated with a hapten, it will generally be desirable for the conjugate to have a molecular weight in excess of 6,000 daltons, although somewhat lower molecular weights

may be useful. Preferably, the natural carrier has some solubility in water or aqueous alcohol. Desirably, the carriers are nontoxic to the animals to be used for generating antibodies.

- The carrier can be coupled to antigens by any available means, included
- 5 the procedures provided herein. Preferably, a carrier moiety has a plurality of hapten or antigen moieties coupled to it, for example, about 15 to 30 for a protein of 100,000 daltons. While steric hindrance and reduced structural complexity may reduce the number of haptens or antigens attached to the carrier, the maximum number is preferred. For example, up to about 25 to about 50
 - 10 hapten moieties can be coupled to BSA carriers.

The use of monoclonal antibodies as biosensors of this invention is preferred because monoclonal antibodies are homogeneous and can be continuously produced in large quantities. Monoclonal antibodies are prepared by recovering lymph node or spleen cells from immunized animals and

- 15 immortalizing the cells in conventional fashion, e.g., by fusion with myeloma cells or by Epstein-Barr virus transformation. Clones expressing the desired antibody are identified by screening cell line media for reactivity with the antigen used to immunize the animals. One of skill in the art can use readily available methods to make monoclonal antibodies, for example, by using the
- 20 hybridoma technique described originally by Kohler and Milstein, Eur. J. Immunol., 6:511 (1976), by Hammerling et al., in "Monoclonal Antibodies and T-Cell Hybridomas", Elsevier, New York, pp. 563-681 (1981), and by Zola, in "Monoclonal Antibodies: A Manual of Techniques", CRC Press, Boca Raton, Fla. (1987). The hybrid cell lines can be maintained in vitro in cell culture
- 25 media. The cell lines producing the antibodies by these procedures can be selected and/or maintained in a medium containing hypoxanthine-aminopterin thymidine (HAT). However, once the hybridoma cell line is established, it can be maintained on a variety of nutritionally adequate media. Moreover, the hybrid cell lines can be stored and preserved in any number of conventional ways,
- 30 including freezing and storage under liquid nitrogen. Frozen cell lines can be revived and cultured indefinitely with resumed synthesis and secretion of monoclonal antibody.

The secreted antibody is recovered from tissue culture supernatant or ascites fluid by conventional methods such as immune precipitation, ion exchange chromatography, affinity chromatography such as protein A/protein G column chromatography, or the like. The antibodies described herein are also

5 recovered from hybridoma cell cultures by conventional methods such as precipitation with 50% ammonium sulfate. If desired, the purified antibodies can then be sterile filtered before use.

The term "monoclonal antibody" as used herein refers to any antibody obtained from a population of substantially homogeneous antibodies, i.e., the

10 individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts.

Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to polyclonal antibody preparations, which typically include different antibodies directed against different

15 determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they are synthesized by the hybridoma culture, uncontaminated by other immunoglobulins.

The monoclonal antibodies herein also include hybrid and recombinant

20 antibodies produced by splicing a variable (including hypervariable) domain of an anti-adduct antibody with a constant domain (e.g. "humanized" antibodies), or a light chain with a heavy chain, or a chain from one species with a chain from another species, or fusions with heterologous proteins, regardless of species of origin or immunoglobulin class or subclass designation, as well as antibody

25 fragments (e.g., Fab, F(ab')₂ and Fv), so long as they exhibit the desired biological activity. See e.g. Cabilly et al. U.S. Pat. No. 4,816,567; Mage and Lamoyi, "Monoclonal Antibody Production Technique and Applications", pp. 79-97 (Marcel Dekker, Inc., New York, 1987).

Thus, the modifier "monoclonal" indicates the character of the antibody

30 as being obtained from a substantially homogenous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies to be used in accordance with

the present invention may be made by the hybridoma method described by Koehler and Milstein, *supra*, or may be made by recombinant DNA methods (Cabilly, et al. *supra*).

In one embodiment, the biosensor antibodies or the present invention are 5 used for diagnostic or imaging purposes *in vivo*, within a mammalian subject. While the *in vivo* use of a monoclonal antibody from a foreign donor species in a different host recipient species is usually uncomplicated, an adverse immunological response by the host to antigenic determinants present on the donor antibody can sometimes arise. In some instances, this adverse response 10 can be so severe as to curtail the *in vivo* use of the donor antibody in the host. Further, the adverse host response may serve to hinder the intercellular adhesion-suppressing efficacy of the donor antibody. Methods to avoid such adverse reactions are available. For example, humanized antibodies or chimeric 15 antibodies (Sun, et al., *Hybridoma*, 5 (Supplement 1):S17, 1986; Oi, et al., *Bio Techniques*, 4(3):214, 1986) can be used. Chimeric antibodies are antibodies in which the various domains of the antibodies' heavy and light chains are coded for by DNA from more than one species. Typically, a chimeric antibody will 20 comprise the variable domains of the heavy (V_H) and light (V_L) chains derived from the donor species producing the antibody of desired antigenic specificity, and the constant domains of the heavy (C_H) and light (C_L) chains derived from the host recipient species. It is believed that by reducing the exposure of the host 25 immune system to the antigenic determinants of the donor antibody domains, especially those in the C_H region, the possibility of an adverse immunological response occurring in the recipient species will be reduced. Thus, for example, it is possible to produce a chimeric antibody for *in vivo* clinical use in humans which comprises mouse V_H and light V_L domains coded for by DNA isolated from ATCC HB X, and C_H and C_L domains coded for with DNA isolated from a human leukocyte.

The present invention further provides a composition comprising an 30 antibody with a functional molecule attached by the methods of the present invention and a suitable carrier. Further, the present invention also provides a therapeutic composition comprising an effective amount of the antibody-

functional molecule conjugate produced by the present methods and a pharmaceutically acceptable carrier.

As used herein, the term "pharmaceutically acceptable carrier" encompasses any of the standard pharmaceutically accepted carriers, such as phosphate buffered saline solution, water, emulsions such as an oil/water emulsion or a triglyceride emulsion, various types of wetting agents, tablets, coated tablets and capsules. An example of an acceptable triglyceride emulsion useful in intravenous and intraperitoneal administration of the compounds is the triglyceride emulsion commercially known as Intralipid R™. Typically such carriers contain excipients such as starch, milk, sugar, certain types of clay, gelatin, stearic acid, talc, vegetable fats or oils, gums, glycols, or other known excipients. Such carriers may also include flavor and color additives or other ingredients.

Methods of Using Biosensors

Biosensors of the invention can be used in vitro or vivo. Biosensors can be used with any type of test sample, such as cultured cells, tissue samples, cell suspensions, cell lysates, partially purified isolates of a potential target and purified isolates of a potential target. A sample that includes cells can be in any form convenient for observation of the target, for example, cells plated on a culture dish, cells on a microscope slide, suspensions of cells, tissues in physiological or culture media or tissues on a microscope slide

In general, the biosensor is contacted with a sample that may contain a target of interest under conditions and for a time sufficient to permit interaction or binding of the biosensor to the target of interest. The biosensor can be contacted with sample that is in solution by simply mixing the biosensor into the solution. When the target is in a cell or tissue, the biosensor can be injected into the cell or tissue. In some experiments placing the needle into the region just adjacent to the nucleus produced a good combination of efficient injection and cell health. Alternatively, the cell or tissue can be transfected or transformed with a nucleic acid capable of expressing the biosensor. One of skill in the art can use readily available procedures for injecting and transfecting cells with such biosensors.

Conditions sufficient to permit interaction of the biosensor with target are generally conditions that permit protein-protein interactions. A temperature sufficient to permit protein-protein interactions is a temperature that maintains and encourages secondary and tertiary polypeptide structure formation. For 5 example, the temperature can be about 4°C to about 40°C, and preferably a temperature of about 4°C to about 37°C. When the sample includes cells, the temperature is preferably a temperature which maintains cellular function, for example, a temperature of about 20°C to about 38°C. A time sufficient for interaction of the biosensor with a target can be determined by one of skill in the 10 art. Such a time can be, for example, about 1 minute to about 2 hours, preferably about 3 minutes to about one hour. In one experiment, visualization of biosensor-target interaction was successfully performed after cells were given about 5-10 minutes to recover after injection.

The invention will now be illustrated by the following non-limiting 15 Examples.

Example 1. Site-specific labeling of a secondary amiooxy group.

Under controlled pH conditions, the low pKa and enhanced nucleophilicity of an aminoxy group relative to other nucleophilic side chains found in peptides suggested the possibility of site-specific reaction with standard 20 electrophiles such as succimidyl esters (Figure 1). While selective labeling of a primary aminoxy group in the context of an unprotected peptide was achieved, extensive attempts to utilize the primary aminoxy group during synthesis failed. Even when protected as the 2-chlorobenzylloxycarbonyl carbamate, deprotonation of the primary aminoxy group allowed rapid acylation, so it 25 could not be readily incorporated during peptide synthesis. Thus, under certain conditions, the use of a secondary aminoxy group may be preferred.

A test peptide containing both a secondary aminoxy group and nucleophilic amino acids that were most likely to interfere with selective labeling at the aminoxy nitrogen (lysine, cysteine, and the amino-terminus) was 30 prepared. As illustrated in Figure 2, NH₂-AKAARAAAK*AA[RACA-CO₂H (SEQ ID NO: 2), here designated SA-test peptide, was synthesized by incorporation and deprotection of N-(2-Cl-benzylloxycarbonyl)-N-

methylaminoxy acetic acid (Figure 4, 3) during solid phase peptide synthesis. The reactivity of the protected secondary amine group was sufficiently attenuated to remain unreactive during Boc solid-phase peptide synthesis on thioester-linker resins. The 2-Cl-Z protection for the N-methylaminoxy amino acid was efficiently removed by standard HF cleavage procedures.

Conditions for selectively labeling the secondary amine group were determined by varying the reaction pH and dye stoichiometry. Labeling with the succinimide ester of tetramethylrhodamine (TMR-OSu) was determined to be optimal in a solvent system consisting of 50%DMSO/50% aqueous acetate buffer at pH of 4.7 with 2 equivalents of dye per mole of peptide (Table 1). The crude reaction products were separated from unreacted dye, and characterized by RP-HPLC and ESI-MS. Under these conditions, a single molecule of dye was incorporated on the amine group with a 78% yield, based on HPLC quantification. However, side reaction products were also isolated and determined to be either SA-test peptide labeled with two dye molecules (~13%) or acetylated peptide products (~5%). The selectivity, as defined by the ratio of the peak areas of desired single-labeled product over double-labeled products, was 6/1 (Table 1).

20

Table 1: Labeling of SA-Test Peptide

Reaction	Buffers		Dye		Percentage			Selectivity
	Gel	TCEP	pH*	Equivalents	SM	1Dye	2Dye	
	Filtration			Lents				
Acetate	(NH ₄)HCO ₃	No	4.7	1.0	32	56	7	8:1
Acetate	(NH ₄)HCO ₃	No	4.7	1.5	14	63	13	4.5:1
Acetate	(NH ₄)HCO ₃	No	4.7	2.0	0	78	13	6:1
Acetate	(NH ₄)HCO ₃	No	4.7	2.4	0	76	17	4.5:1
Acetate	(NH ₄)HCO ₃	No	4.7	3.0	0	71	23	3:1
Citrate	0.1% TFA	Yes	4.7	4.3	4.3	89.6	4.1	22:1
Carbonate	0.1% TGA	Yes	9.0	0.99	13	85	0	51:1*

*

It was determined that many of the side reactions occurred during size exclusion chromatography in the ammonium bicarbonate solvent used to separate reaction products. Using an acidic solvent system, 0.1% TFA, virtually eliminated multiply labeled side products leading to considerable improvements in both selectivity and yield. Including a mild reducing agent, tris(2-carboxyethyl)phosphine (TCEP), in the reaction buffer also significantly curtailed several minor side reactions revealed by HPLC, especially disulfide formation. Labeling and gel filtration under these optimized conditions produced a 70% recovered yield of labeled SA-test peptide (90% yield based on HPLC quantitation), 90% of which was labeled with only a single dye at the aminoxy amine. The labeling selectivity was increased to 22/1 (Table 1).

The purity of the product was confirmed using RP-HPLC, mass spectroscopy, further chemical reaction, and isolation of purposefully overlabeled products. The labeled peptide eluted as a single peak under all HPLC conditions tested with a mass consistent with that predicted for singly-labeled SA-test peptide (Mass=1970). To determine that the labeling site was indeed at the N-methylaminoxy group, a selective zinc/acetic acid reduction procedure was used to cleave the N-O bond (Figure 3). HPLC of the reduction reaction showed >98% conversion of the starting material and a new earlier eluting peak. The mass of this peak (1530 amu) corresponded to the predicted mass of the unlabeled SA-test peptide cleaved at the aminoxy N-O bond. The residual zinc was washed several times with a saturated solution of EDTA in water, which demonstrated that the reduction reaction was complete.

To eliminate the unlikely possibility that the HPLC peak containing isolated single-labeled SA-test peptide product was a mixture of two labeled species, SA-test peptide was reacted under higher pH conditions (pH 9.0) to label all reactive sites. SA-test peptide contained 3 nucleophilic labeling sites which would be irreversibly labeled: aminoxy, lysine, and N-terminal amine. Dye labeling at high pH generated a mixture of peptides labeled at all possible combinations of sites with 1, 2 or 3 dyes. HPLC analysis of this reaction

mixture showed 8 peaks, indicated by ESI-MS to correspond to unreacted SA-test peptide, SA-test peptide single-labeled at the aminoxy nitrogen, and six additional peaks corresponding to two single-labeled peptides, three peptides bearing two dyes and a single triply-labeled peptide species. This experiment
5 revealed the HPLC retention times of all these products, none of which co-eluted with the peak identified as SA-test peptide labeled with a single dye on the secondary aminoxy nitrogen.

Site-specific labeling of the aminoxy group could even be achieved at basic pH (Table 1). Using pH 9.0 carbonate buffer in our solvent system,
10 addition of 0.5 equivalents of dye produced, after 3 hours, ~50% conversion of the starting SA-test peptide to a single peak with the elution time of the desired singly-labeled product. After addition of another 0.5 equivalents of TMR-OSu and an additional 3 hours of reaction time, HPLC showed ~85% conversion to a peak with the retention time of the desired product. Two minor peaks (~2-3% of
15 total peak area), were also apparent and corresponded to the two other single-labeled SA-test peptide species identified in the multiple labeling experiment above. The N-hydroxysuccinimide ester of rhodamine clearly showed selective reactivity with the aminoxy group.

The selectivity observed at higher pH cannot be explained by the
20 nucleophilicity of the aminoxy group alone. In fact, others have shown that at, in an uncatalyzed reaction with phenyl acetate at high pH, amines are more reactive than O-alkyl aminoxy groups. Therefore we suggest that kinetic factors are contributing to the selective reactivity of the N-methylaminoxy group, even when competing groups are not protonated. Possible reasons for this
25 include: (1) the aminoxy oxygen localizes the nitrogen near the activated ester via formation of a hydrogen bonded “bridged” intermediate (2) a base catalyzed reaction pathway under the conditions of our reaction. This exceptional reactivity has important practical implications, as it can allow the selective labeling of acid-labile polypeptides and synthetic proteins under physiological or
30 basic conditions.

Example 2. The secondary aminoxy group is compatible with

C^a-thioesters and amide-forming ligation.

Preparation of proteins by total chemical synthesis often requires the ligation of large polypeptides prepared by solid-phase peptide synthesis on thioester-linker resins. The most generally applicable methods available for 5 ligations are native chemical ligation and expressed protein ligation. These processes utilize the same basic chemistry to join two peptides, one with an N-terminal cysteine and the other with a C-terminal thioester, through a regiospecific and site-specific reaction to generate a larger polypeptide. The application of aminoxy-labeling chemistry to the synthesis of large 10 polypeptides and proteins requires compatibility with these solid-phase peptide synthesis and ligation chemistries.

The optimal approach for utilizing aminoxy-labeling chemistry in the chemical synthesis of proteins is direct incorporation of the aminoxy group as part of an amino acid used during standard solid-phase peptide synthesis. For 15 this purpose, we generated a suitably protected N-methylaminoxy amino acid, α-Boc-β-[N-(2-Chlorobenzyloxycarbonyl)-N-Methylaminoxy Acetyl]-α,β-Diaminopropionic Acid [Boc-2-Cl-Z-(SA)Dapa-OH] (4), as shown in Figure 4. This amino acid, referred to as SAOD, was incorporated into the peptide sequence LY-(SAOD)-AG-MPAL thioester by synthesis on TAMPAL 20 thioester-linker resin, as described below in the Methods. (MPAL is the C-terminal mercaptopropionyl-leucine group generated by cleavage of a peptide from TAMPAL resin, see Hojo, H., et al., *Bull. Chem. Soc. Jpn.* 1993, 66:2700-2706; and Hackeng, T.M. et al., *Proc. Natl. Acad. Sci. USA.* In press)

Ligation of the LY-(SAOD)-AG-MPAL (SEQ ID NO: 3) thioester 25 peptide with the peptide CRANK-NH₂ (SEQ ID NO: 4), was tested using standard procedures employing phosphate buffer with 6M guanidine hydrochloride at neutral pH in the presence of 2-3% thiophenol by volume. The ligation proceeded over 24 hours and generated the desired ligation product, LY-(SAOD)-AGCRANK-NH₂ (SEQ ID NO: 5), at ~85% yield. The major side 30 product was attributable to modification of unligated CRANK-NH₂ peptide under the ligation conditions (mass=714.5, data not shown), and was not related to the presence of the aminoxy group. There was also a single time-dependent

side reaction, which generated a product of 14 mass units lower than the desired. Using high concentrations of reacting peptides and isolating the ligation product after 24 hours reduced this side reaction to acceptable levels (<5%).

The ligation of a peptide containing multiple potentially reactive

- 5 functional groups, including a hexahistidine tag useful for affinity chromatography was also tested. Coupling CEYRIDRVRLFVDKLDNIAQVPRVGAA-HHHHHH (SEQ ID NO: 6) to LY-(SAOD)-AG-MPAL thioester proceeded to completion in 5 hours with minimal side reactions. In both ligation reactions, there was less than 1% LY-(SAOD)-
10 AG-MPAL self condensation product, indicating that the aminoxy group and thioester do not appreciably react with one another under the ligation conditions. These results demonstrate that the inclusion of an unprotected aminoxy group in the peptide chain is compatible with native chemical ligation.

Labeling of the two ligation products using tetramethylrhodamine

- 15 succinimide ester proceeded with selectivity similar to that for the SA-test peptide. HPLC integration indicated that the product of the LY-(SAOD)-AG-MPAL ligation with CRANK-NH₂ was labeled with greater than 95% efficiency and with a selectivity of 34:1. Mass spectral analysis and zinc reduction demonstrated labeling at only the aminoxy group. For the longer hexahistidine-
20 containing polypeptide ligation product, selectivity for the aminoxy group was greater than 10:1, but it was difficult to achieve high yields. The histidines could potentially have been affecting yield and selectivity by catalyzing nucleophilic attack on the succinimide ester of the reactive dye. Inclusion of guanidine hydrochloride in the reaction solvent increased the yield to approximately 50%,
25 indicating that folding or poor solubility of the peptide was a factor in preventing access of the reactive dye to the aminoxy group. Selectivity was also improved, presumably because of the availability of the reactive secondary aminoxy group. Single-site labeling at the aminoxy group was proven by mass spectral analysis of trypsin and α -chymotrypsin digests of the labeled
30 polypeptide product.

Example 3. Specificity of labeling in protein domains containing aminoxy amino acids.

As a control to establish the selectivity of labeling for aminoxy amino acid, the labeling of native β-Lactoglobulin with tetramethylrhodamine

- 5 N-hydroxysuccinimide ester was attempted. It was found that non-specific labeling of this 162 aa protein containing 15 lysines and 4 cysteines was minimal (<1%), even after 6 hours.

Finally, the GTPase binding domain of p21 activated kinase (45 aa, 4 lys, 1 cys) with a secondary aminoxy amino acid incorporated at the amino-

- 10 terminus (SAOD-PBD) was prepared. Previous experiments have demonstrated that PBD domains labeled with fluorescent reporter dyes at this terminus could be used as biosensors of GTPase activation. Labeling of PBD using the new methodology would enable the production of sufficient quantities to apply the biosensors *in vivo* and in pharmaceutical screening applications, and would 15 allow incorporation of sensitive detectable groups enabling applications within living cells.

SAOD-PBD was readily labeled with Alexa-532 N-hydroxy-succinimide ester by titration addition of dye at pH 4.7 over 72 hours. The labeling efficiency was commensurate with that reported for the longest model peptide 20 (~50% yield by HPLC quantitation) and there was no indication of multiple labeling. In this case, isolation of labeled SAOD-PBD by RP-HPLC proved difficult. Separation to baseline resolution was not achieved, but small quantities of unlabelled PBD in the labeled product do not preclude the use of the labeled material in biosensor applications. Previous reports indicate that 25 separation of labeled product from starting polypeptide is highly dependent on the specific peptide and the attached dye.

These results demonstrate that the optimized site-specific labeling chemistry reported here is compatible with the steps required for the preparation of proteins by total chemical synthesis.

- 30 **Example 4. Materials and Methods.**

General: For column chromatography, silica gel (230-400 mesh) was used in standard glass columns with gravity or air pressure. Reversed-phase high

performance liquid chromatography (RP-HPLC) was performed on a Waters HPLC system with UV detection at 214 nm using either a Vydac C-18 analytical column (5 μ m, 0.46 x 25cm), a Waters RCM 8 x 10 module equipped with a semi-preparative Delta Pack C-18 Radial Pack cartridge column (15 μ m, 8 x 100 mm) from Millipore, or a Vydac C-18 preparative scale column (15 μ m, 1.0 x 25 cm). Linear gradients of solvent B (0.09% TFA in 90% acetonitrile/10% water) in solvent A (0.1% TFA in water) were used for all HPLC chromatographic separations.

Mass spectra of peptides were obtained either with a Sciex API-III 10 electrospray ionization (ESI) triple-quadrupole mass spectrometer (PE Biosystems, Foster City), or matrix assisted laser desorption ionization time-of-flight (MALDI-TOF) instruments from Thermo Bioanalysis (Thermo Bioanalysis, LTD., UK) or Kratos Analytical (Chestnut Ridge, NY). For ESI-MS, the observed masses reported were derived from the experimental m/z states 15 for all observed charge states of a molecular species using the program MacSpec (Sciex, version 2.4.1) for electrospray mass spectrometry. MALDI-MS observed masses were relative to internal calibration using α -cyano-hydroxycinnamic acid or sinipinic acid matrices. Calculated masses reported were derived from either MacProMass (Terry Lee and Sunil Vemuri, Beckman Research Institute, 20 Duarte, CA) or PAWS (Version 8.1.1, ProteoMetrics) and reflect the average isotope composition of the singly-charged molecular ion. Proton nuclear magnetic resonance spectrometry was recorded on a Bruker AC-250 mass spectrometer and data was analyzed using WinNMR (Bruker Instruments). Ultraviolet-Visible spectroscopy was performed on a Hewlett-Packard 25 photodiode-array spectrophotometer.

Boc-L-amino acids were purchased from Novabiochem (La Jolla, CA) or Bachem Bioscience, Inc. (King of Prussia, PA). [[4-(Hydroxymethyl)phenyl]-acetamido]methyl (-OCH₂-Pam) Resin was purchased from PE Biosystems (Foster City, CA) and methylbenzhydrylamine (MBHA) resin was purchased 30 from Peninsula Laboratories, Inc. (San Carlos, CA). Solvents were Synthesis grade or better and were purchased from Fisher Scientific (Tustin, CA). Trifluoroacetic acid (TFA) and anhydrous hydrogen fluoride were purchased

from Halocarbon (New Jersey) and Matheson Gas (Rancho Cucamonga, CA). Dyes were obtained from Molecular Probes (Eugene, OR). All other reagents were analytical grade or better and were purchased from Aldrich (Milwaukee, WI), Lancaster (Windham, NH), Peptides International (Louisville, KY) or 5 Richelieu Biotechnologies (Montreal, Canada).

- Peptide Segment Synthesis.** Synthesis of peptides was carried out manually using optimized stepwise solid-phase synthesis methods with *in situ* neutralization and HBTU activation procedures for Boc chemistry on either -
- 10 OCH₂-Pam, MBHA, or Trt-protected mercaptopropionyl-Leu (TAMPAL) resin (Hojo, H., et al., *Bull. Chem. Soc. Jpn.* 1993, 66:2700-2706; Hackeng, T.M. et al., *Proc. Natl. Acad. Sci. USA*. In press; and Schnolzer, M., et al., *Int. J. Peptide Protein Res.* 1992, 40:180-193). Standard Boc protecting group strategies were employed. Coupling was monitored by quantitative ninhydrin 15 assay after 15 minute coupling cycles. After chain assembly, standard deprotection and cleavage from the resin support was carried out by treatment at 0 °C for 1 hour with anhydrous HF containing either 10% p-cresol or anisole as scavenger. Purification was performed using RP-HPLC.
- 20 **Synthesis of TAMPAL Resin** (Hojo, H., et al., *Bull. Chem. Soc. Jpn.* 1993, 66:2700-2706). 2.5 grams of MBHA resin (0.865 mmol/g, 2.16 mmol of amine) was swelled in DMF. Boc-Leu-OH (1.1 grams, 4.4 mmol) was activated with HBTU (8 ml, 0.5M solution) and DIEA (2 ml), then coupled to the MBHA resin until complete reaction by ninhydrin assay. The N^α-Boc group of the 25 linked leucine was removed with neat TFA, then S-Trt-β-mercaptopropionic acid (1.5 grams, 4.3 mmol), activated in the same manner as Boc-Leu-OH, was added to the deprotected Leu-MBHA resin and allowed to couple until complete reaction. The S-Trt-β-mercaptopropionyl-Leu-MBHA resin was washed extensively with DMF, then DCM/MeOH (1/1), and finally dried *in vacuo* to 30 yield 3.39 grams of thioester resin. Substitution calculated by weight gain yielded 0.549 mmol/gram.

Deprotection of TAMPAL Resin: S-trityl protection was removed by two 5 minute treatments with 95% TFA/5% triisopropylsilane. The deprotected resin was extensively washed with DMF before coupling the first amino acid, activated using optimized *in situ* neutralization protocols.

5

Synthesis of N-(2-Cl-benzyloxycarbonyl)-N-Methylhydroxylamine (1)

(Jencks, W.P., Carriuolo, J. *J. Am. Chem. Soc.* 1960, 82:675; Jencks, W.P. *J.*

Am. Chem. Soc. 1958, 80:4581, 4585). N-methylhydroxylamine hydrochloride

(0.95g, 11.37mmol) was dissolved in 3ml of water with rapid stirring. The pH

- 10 of this solution was adjusted to 6-7 by dropwise addition of a saturated solution of sodium bicarbonate. 2-Chlorobenzyloxycarbonyl-N-hydroxysuccinimidyl carbonate (1.2g, 4.23mmol) was dissolved in 4ml of THF and added slowly to the rapidly stirring solution of neutralized N-methylhydroxylamine. After stirring at room temperature for 14 hours, the reaction was quenched with 20ml
15 of saturated sodium bicarbonate and extracted three times with 20ml ethyl acetate. The combined ethyl acetate layers were washed once with saturated sodium bicarbonate, dried over anhydrous sodium sulfate and the solvent was removed *in vacuo* to yield 0.77g (3.77mmol, 84%) of an off-white solid. TLC R_f=0.2 (Hex/EtOAc/AcOH 80/20/1). ¹H NMR: 3.23(s, 3H), 5.25 (s, 2H), 7.24
20 (m, 2H), 7.37 (m, 2H). HRMS: Expected=216.0427, Observed=216.0425.

Synthesis of N-(2-Cl-benzyloxycarbonyl)-N-Methylaminoxy Acetic Acid-

Tert-butyl ester (2) (Jerry March *in* Advanced Organic Chemistry, Third

Edition. John Wiley & Sons, New York. 1989, pp381; and Nyberg, D.D.,

- 25 Christensen, B.E. *J. Am. Chem. Soc.* 1957, 79:1222; Motorina, I.A., et al., *Synlett* 1996, 389). Compound 1 (0.96g, 4.71mmol) was dissolved at room temperature in 10ml of THF with rapid stirring. Bromoacetate tert-butyl ester (1.05g, 5.38mmol) was added, then sodium iodide (1.5g, 10.01mmol) followed by DIEA (2.5ml, 15.92mmol). The reaction changed to an orange-yellow color
30 after addition of sodium iodide. The reaction was quenched with 30ml water after complete reaction (~3 hours) and extracted 3 times with ethyl acetate. The combined ethyl acetate layers were dried over sodium sulfate and the ethyl

acetate was removed *in vacuo*. The resultant oily solid was purified by silica chromatography on 230-400 mesh silica gel using hexanes/ethyl acetate/acetic acid (80/20/1) to yield 1.40g (4.29mmol, 90%) of a pure yellow oil. TLC R_f=0.5 (Hex/EtOAc/AcOH 80/20/1). ¹H NMR: 1.46 (s, 9H), 3.29 (s, 3H), 4.36 (s, 2H), 5 5.26 (s, 2H), 7.24 (m, 2H), 7.40 (m, 2H). HRMS: Expected Mass=330.1108, Observed Mass=330.1104.

Synthesis of N-(2-Cl-benzyloxycarbonyl)-N-Methylaminoxy Acetic Acid (3) (Bryan, D.B., et al., *J. Am. Chem. Soc.* 1977, 99:2353). Compound 2 (1.1g, 10 3.30mmol) was dissolved into 4ml of DCM and, with rapid stirring, neat TFA (5ml) was added dropwise over 2 minutes at room temperature. After 1 hour, the reaction was quenched with 20ml water, extracted 3 times with DCM, and the combined DCM layers were dried over sodium sulfate. The DCM was removed *in vacuo* to yield 0.9g (3.28mmol, 99%) of an off-white solid. TLC R_f=0.2 (Hex/EtOAc/AcOH 80/20/1). ¹H NMR: 3.23 (s, 3H), 4.50 (s, 2H), 5.32 (s, 2H), 15 7.28 (m, 2H), 7.40 (m, 2H). HRMS: Expected Mass=274.0482, Observed Mass=274.0479.

Synthesis of N-(2-Cl-benzyloxycarbonyl)-N-Methylaminoxyacetyl- α -Boc-
20 **α,β -Diaminopropionic Acid [(SA)Dapa-OH] (4)** (Wahl, F., Mutter, M. *Tett. Lett.* 1996, 37:6861-6864; and Anderson, G.W., et al., *J. Am. Chem. Soc.* 1964, 86:1839). N-(2-Chlorobenzyloxycarbonyl)-N-Methylaminoxy Acetic Acid (3) (2.5g, 9.2mmol) was activated with N-hydroxysuccinimide (2.11g, 2equiv.) and DIC (1.440ml, 1.0equiv.) in 20ml DCM. This reaction was rapidly stirred at
25 room temperature for 2 hours prior to the addition of N ^{α} -Boc- α,β -diaminopropionic acid (2.3g, 1.2equiv.) and DIEA (3.20ml, 2equiv.). After 4 hours, the DCM solvent was removed *in vacuo*, and 50ml ethyl acetate was added. The ethyl acetate layer was washed twice with 0.5M acetate buffer, pH=4.0, then twice with 0.1N sulfuric acid. The combined acid washes were
30 then washed with 50ml ethyl acetate. The combined ethyl acetate layers were dried over sodium sulfate, then concentrated *in vacuo* to yield a viscous yellow oily solid. This solid was subjected to 3 hexane precipitations from diethyl ether

to yield 2.16g (51% yield) of an off-white solid. TLC R_f=0.2-0.4
(Hex/EtOAc/AcOH 30/70/0.5). ¹H NMR: 1.45 (s, 9H), 3.16 (s, 3H), 3.54 (d-of-t, 1H, J=14.3, 4.6Hz), 3.93 (m, 1H, J=14.3, 7.5, 4.6Hz), 4.31 (s, 0.5H), 4.38 (s, 1H), 4.45 (s, 1H), 4.51 (s, 0.5H), 5.34 (s, 2H), 5.97 (broad-d, 1H, J=7.3Hz), 7.30
5 (m, 2H), 7.43 (m, 2H), 8.50 (broad-s, 1H). HRMS: Expected Mass=460.1487, Observed Mass=460.1480.

Synthesis of Secondary Aminooxy Test Peptide (SA-test peptide). The SA-test peptide, NH₂-AKAARAAAAK*AARACA-CO₂H, was synthesized with
10 Lys 10 side chain Fmoc protection as described previously (Canne, L.E., et al., *J. Am. Chem. Soc.* 1995, 117:2998-3007). Incorporation of the secondary aminooxy group was accomplished by coupling 2-Cl-Z protected N-methylaminooxyacetic (300mgs, 1.09mmol) activated with Diisopropylcarbodiimide (157ul, 1.00mmol) and N-hydroxysuccinimide (140mgs, 1.22mmol) in 2 ml DCM for 1-2 hours, then diluted with 2 ml DMF just prior to coupling to the ε-amino group of Lys 10. Optimized coupling, cleavage and purification protocols were utilized. Amino acid analysis was consistent with the desired peptide. Expected Mass=1560, Observed Mass=1559.

20

Synthesis of LY-(SAOD)-AG-MPAL-Thioester. LY-(SAOD)-AG-MPAL-Thioester was synthesized using optimized *in situ* neutralization protocols for Boc chemistry on TAMPAL resin. Coupling of the N^α-Boc-(SA)Dapa-OH amino acid was accomplished by reacting the *in situ* activated N-hydroxysuccinimide ester to the deprotected amino-terminal nitrogen of alanine (Canne, L.E., et al., *J. Am. Chem. Soc.* 1995, 117:2998-3007). (SA)Dapa-OH (4) (230mgs, 0.5mmol) was dissolved in 1ml DCM and N-hydroxysuccinimide (115.1mgs, 1.0mmol) and DIC (74.4μl, 0.47mmol) were added. The reaction was mixed briefly and allowed to activate for 1-2 hours at room temperature
25 prior to coupling to the deprotected N-terminus of the peptide chain. After this coupling, no further modifications of the synthetic protocols were required.
30 Expected mass=797, Observed mass=797.

Ligation of LY-(SAOD)-AG-MPAL-Thioester with CRANK-NH₂ Peptide.

LY-(SAOD)-AG-MPAL-Thioester (3mg, 3.8μmol) was dissolved into 100 ul of 50mM phosphate buffer containing 6M guanidine hydrochloride, pH=7.2. To 5 this solution was added CRANK-NH₂ peptide dissolved into 100μl of the same phosphate buffer and 3ul of thiophenol. The reaction was monitored by analytical reversed-phase HPLC. After 24 hours, the ligated product, LY-(SAOD)-AGCRANK-NH₂ (SEQ ID NO: 10), was isolated by semi-preparative reversed-phase HPLC (gradient=10-50% B over 60 minutes) and lyophilized to 10 yield a fluffy white solid. Amino acid analysis was consistent with the desired product peptide. Expected Mass=1168, Observed Mass=1168.

**Ligation of LY-(SAOD)-AG-MPAL-Thioester with
CEYRIDRVRLFVDKLDNIAQ-VPRVGAA-HHHHHH (SEQ ID NO: 7).**

15 LY-(SAOD)-AG-MPAL (0.3mgs, 0.37mmol) and CEYRIDR-VRLPFVDKLDNIAQ-VPRVGAA-HHHHHH (1.5mgs, 3.8mmol) were subjected to the same ligation and purification conditions as described above to yield 1.0mgs (58% yield) of a white fluffy solid. Expected Mass=4518, Observed Mass=4517.

20

Synthesis of Amino-Terminal P21 Binding Domain (PBD) Peptide Fragment, (SAOD)-KKKEKERPEISLPSDFEHTIHVGFDA-MPAL Thioester (SEQ ID NO: 8): Secondary amiooxy containing amino-terminal PBD thioester were synthesized as described above using TAMPAL resin. HF 25 cleavage utilizing p-Cresol scavenger followed by HPLC purification yielded (SAOD)-KKKEKERPEISLPSDFEHTIHVGFDA-MPAL containing two DNP groups protecting the histidines. Mass Expected=3745, Mass Observed=3745.

Synthesis of Carboxy-Terminal of P21 Binding Domain (PBD) Peptide Fragment, CTGEFTGMPEQWARLLQT (SEQ ID NO: 9): The native carboxy-terminal half of PBD was synthesized using standard FMOC synthesis

protocols by the Scripps Peptide and Protein Core Facility. Mass Expected=2068, Mass Observed=2068.

Synthesis of SAOD-Modified PBD, SAOD-KKKEKERPEISLPSDFEHTIH-

- 5 **VGFDACTGEFTGMPEQWARLLQT (SEQ ID NO: 11):** 1.5 mg of SAOD-KKKEKERPEISLPSDFEHTIHVGFDAMPAL (0.4 mmol) was ligated to 1 mg (4.8 mmol) of carboxy-terminal fragment, CTGEFTGMPEQWARLLQT, as described above. After 48 hours, the ligated PBD proteins were isolated by RP-HPLC and lyophilized, 1.5mg (70% yield), Mass Expected=5262, Mass
10 Observed=5262.

Selective Labeling of SA-Test Peptide with Tetramethylrhodamine N-hydroxysuccinimidyl ester. A solution of SA-test peptide (3.396 μ g/ μ l, 2.18mM) in 5% acetate buffer, pH=4.7 incorporating 5mM TCEP was utilized

- 15 for labeling. A stock solution of dye (5 μ g/ μ l, 9.5 mM) was made by dissolving TMR-OSu in neat DMSO. For each reaction, the dye stock was diluted so that the desired number of dye equivalents could be added in 20 μ l of DMSO. The following equivalents of dye were tested: 1.2, 1.5, 1.8, 2.0, 2.4, 3.0, and 4.3. With constant stirring, 20 μ l of dye solution was added in two 10 μ l aliquots to
20 20 μ l of peptide solution at room temperature. The second aliquot of dye in DMSO was added 10 minutes after the initial dye addition. After complete addition of dye, the reaction was briefly vortexed, then incubated at room temperature. After 3 hours, the labeled reaction product(s) were separated from unreacted dye by gel filtration on Sephadex G-10 or G-15 columns using either
25 100 μ M ammonium bicarbonate or, after optimization, 0.1% TFA in water. The individual peptide product(s) were then separated by RP-HPLC and analyzed by ESI-MS. Mass Expected=1971, Mass Observed=1970.

Non-Selective Labeling of SA-Test Peptide with Tetramethylrhodamine N-

- 30 **hydroxysuccinimide ester.** SA-test peptide (3.396 μ g/ μ l, 2.18mM) was dissolved 100mM sodium carbonate, pH=9.01 containing 5 mM TCEP. 18 μ l of neat DMSO was added to 20 μ l of peptide solution. With rapid mixing, 1 μ l stock

dye solution in DMSO was added to this mixture. Upon completion of addition, the reaction was vortexed briefly, then incubated at room temperature. After 3 hours, a 15 μ l aliquot was removed and evaluated by RP-HPLC and ESI-MS. This process was repeated until significant levels of reaction were detected by 5 formation of single-labeled SA-lysine test peptide products.

Selective Labeling of LY-(SAOD)-AGCRANK-NH₂ and LY-(SAOD)-AGCEYRIDRVR-LFVDKLDNIAQVPRVGAA-HHHHHH with Tetramethylrhodamine N-hydroxy-succinimidyl Ester. A sample of 20 μ L 10 of LY-(SAOD)-AGCRANK-NH₂ (2.5 μ g/ μ l, 2.18mM) in 200mM citrate buffer, pH=4.7, 5mM TCEP was labeled using 4.3 equivalents of dye in DMSO, purified and analyzed. Expected Mass=1580, Observed Mass=1580. A 20 μ L sample of LY-(SAOD)-AGCEYRIDRVRLFVDKLDNIAQVPRVGAA- 15 HHHHHH (SEQ ID NO: 12) (9.7 μ g/ μ l, 2.12mM) in 200mM citrate buffer, pH=4.7, containing 5mM TCEP and 3M guanidine hydrochloride was labeled 20 using a modified procedure. 18 μ l of a solution of tetramethylrhodamine N-hydroxysuccinimide (10 μ g/ μ l) in DMSO was added in 6 μ l aliquots over 15 minutes with rapid mixing. The reaction was incubated at room temperature for 5 hours prior to gel filtration/RP-HPLC purification and mass spectral analysis. 25 Expected Mass=4930, Observed Mass=4929.

Labeling of β -Lactoglobulin with Tetramethylrhodamine N-hydroxysuccinimide ester: 10 μ L of a 10 mg/ml solution of 25 tetramethylrhodamine N-hydroxysuccinimide ester in DMSO (9.5 equivalents compared to protein), was added to a solution containing 10 μ L of DMSO and 20 μ L of a solution of β -Lactoglobulin (20.3 mg/ml or 1.1 mM) in 2.8 M guanidine hydrochloride with 5 mM TCEP (pH 4.7). After 3 hours, protein was separated from unreacted dye by gel filtration, and labeling was determined by analysis of the dye-to-protein ratio (protein concentration was determined by the 30 method of Waddel and ϵ for tetramethylrhodamine in phosphate buffer, pH=8.0 of 81,000).

Labeling of PBD Proteins with Alexa-532 N-Hydroxysuccinimide Ester:

150 µg of the SAOD-modified PBD protein was dissolved in 105 L of 200 mM sodium citrate buffer, pH=4.8, containing 5 mM TCEP (protein concentration ~0.28 mM). A solution of Alexa-532-OSu in DMSO (dye concentration ~10 mg/ml in DMSO) was titrated into the protein solution in 5 µL aliquots over 72 hours. Four hours after each addition, the extent of labeling was determined by RP-HPLC and MS. Labeling was continued until quantities of double-labeled PBD was obtained (SAOD-modified PBD). Alexa-532 labeled SAOD-modified PBD, Mass Expected=5871, Mass Observed=5870.

Zinc/Acetic Acid Reduction of the N-methylaminoxy N-O Bond in Peptides. Reductive cleavage of the N-O bond was performed using zinc and aqueous acetic acid. Effervescence in the reaction was evident after a few seconds and subsided after 60-120 minutes. After 14 hours, the reaction supernatant was analyzed by RP-HPLC and ESI-MS. Reduction of labeled SA-Test peptide, Expected Mass=1530, Observed Mass=1530: Reduction of labeled (SAOD)-AGCRANK-NH₂: Expected Mass=1139, Observed Mass=1139.

20

Trypsin/Chymotrypsin Cleavage of Labeled LY-(SAOD)-AGCEYRIDRVRLFVDKLDNIAQVPRVGAHHHHHH Peptide. 10 µl of a 0.05 mg/ml solution of either trypsin or α-chymotrypsin in 25 mM ammonium carbonate (without pH adjustment) was added to 5 µl of a 10-20 µg/µl solution of pure tetramethylrhodamine-labeled peptide in water (final concentration of protease is 0.033 mg/ml). The reaction was incubated at room temperature for 24 hours prior to analysis of peptide fragments by MALDI-MS.

30 **Example 5: Synthesis and Characterization of Fluorescent Dyes**

Materials and Methods. Commercially available solvents and reagents were purchased from major suppliers. ¹H NMR spectra of dye solutions in DMSO-d₆

were recorded at 500.11 Hz on Bruker Avance DRX-500 MHz. The pentet corresponding to the residual protons of DMSO-d₆ (2.49 ppm) was used as internal reference.

The analytical sample of the dyes was purified by HPLC on a Vydac C-5 18 column (no. 218TP152022, 22*250 mm, 15-20 µm, 3mL/min) using acetonitrile-water gradient elution.

Absorbance spectra were recorded on Hewlett-Packard HP8452 diode array spectrophotometer and fluorescence spectra were recorded on SPEX Fluorolog 2 spectrometer.

All reactions were run in a oven-dried round bottom flask, under argon atmosphere, and capped with a rubber septum. 1-Benzothiophen-3(2H)-one 1,1-dioxide was synthesized according the procedure of M. Regitz (Chem. Ber., 1965, 98: 36). 3-(dicyanomethylene)indan-1-one was synthesized from 1,3-indandione by the method of K. A. Bello, L. Cheng and J. Griffiths (Chem. Soc. Perkin Trans. II, 1987, 815-818).

Preparation of 2Z)-2-[(2E)-3-METHOXYPROP-2-ENYLIDENE]-1-BENZOTHIOPHEN-3(2H)-ONE 1,1-DIOXIDE (Compound 1-SO), (2Z)-2-[(2E)-3-METHOXYPROP-2-ENYLIDENE]-3-(DICYANOMETHYLENE)INDAN-1-ONE (Compound 1-CN).

To a solution of corresponding ketone (2.5 mmol) in 20-50 ml of methanol was added trifluoroacetic acid in one aliquot at 80°C followed immediately by the addition of 2 ml of 1,3,3-trimethoxy propene. The reaction mixture turned into a clear dark brown solution and after an interval of 30 sec a yellow solid precipitated out. The precipitate was filtered and dried.

For compound 1-SO, the yield was 60%. ¹H NMR (CDCl₃): δ 4.04 (s, 3H, CH₃O), 6.53 (t, ³J_{H-H} = 12.1Hz, 1H), 7.53 (d, ³J_{H-H} = 11.7 Hz, 1H), 7.74 (d, ³J_{H-H} = 12.8 Hz, 1H), 8.2-8.4 (m, 4 H, Ph).

For compound 1-CN, the yield was 60%. ¹H NMR(CDCl₃): δ 8.67 (dd, 6.2 Hz, 2.9 Hz, 1H), 8.37 (d, 11.7 Hz, 1H), 7.87 (m, 1H), 7.73 (m, 3H), 7.51 (d, 12.05 Hz, 1H), 3.98 (s, 3H).

PREPARATION OF (2Z)-2-[(2E,4E)-4-(3-(3-SULFONATOPROPYL)-1,3-BENZOTHIAZOL-2(3H)-YLIDENE)BUT-2-ENYLIDENE]-1-BENZOTHIOPHEN-3(2H)-ONE 1,1-DIOXIDE (DYE 2).

1.16 g (2.00 mmol) of 3-(2-methyl-1,3-benzothiazol-3-ium-3-yl)propane-

- 5 1-sulfonate and 625 mg (2.50 mmol) of compound 1-SO were dissolved in 65 ml of mixture CH₃OH-CHCl₃ (1:1) and heated at reflux. 328 mg (4.00 mmol) of AcONa in 10 ml was added at once and the reaction mixture was refluxed for additional 30 min. After cooling, the dye was separated by filtration, recrystallized from methanol and dried. The yield was 65%.
- 10 ¹H NMR: (DMSO-d₆) δ 1.85 (p, ³J_{H-H} = 6.2 Hz , 2H, CH₂-CH₂-CH₂), 2.60 (t, ³J_{H-H} = 6.2 Hz , 2H, CH₂-CH₂-SO₃), 4.45(t, ³J_{H-H} = 6.2 Hz , 1H, CH₂-N), 6.87 (d, ³J_{H-H} = 13.2 Hz, 1H), 7.3 - 8.1 (m, 11H).

Table 2. Selected Merocyanine Dyes.

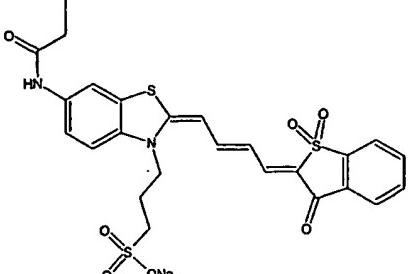
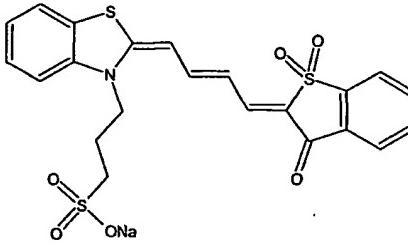
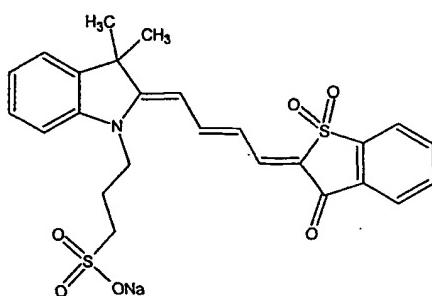
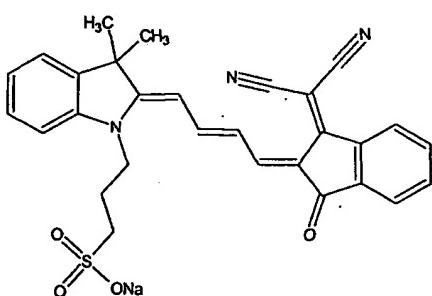
Compound	Dye Structure
1	
2	
3	
4	

Table 3. Spectral Data for Selected Merocyanine Dyes.¹

Compound	Extinction coefficient ($\times 10^{-3}$)	Absorption Maximum	Emission Maximum	Quantum Yield
1	100	618	636	0.07
2	100	603	623	0.10
3	120	586	618	0.17
4	80	667	693	0.10

5 ¹All spectral data are recorded in n-butanol.

Example 6: Imaging the Spatio-temporal Dynamics of Rac Activation *in vivo* with FLAIR

FLAIR (Fluorescent Activation Indicator for Rho GTPases) is a biosensor system that maps the spatial and temporal dynamics of Rac activation in living cells. The approach is based on microinjection of a fluorescently labeled domain from p21-activated kinase into cells expressing GFP-Rac. The injected domain (called PBD, for p21-binding domain) binds only to Rac-GTP, and not to Rac-GDP (Thompson et al., 37 Biochemistry 7885 (1998); Benard et al., 274 J. Biol. Chem. 13198 (1999)). Within living cells, PBD binds to the GTP-Rac wherever it has bound GTP, bringing the Alexa546 dye on the PBD near the GFP on the Rac to produce fluorescence resonance energy transfer (FRET). Thus, the FRET signal marks subcellular locations where Rac is activated. This can be quantified to follow the changing levels and locations of Rac activation or to trace the kinetics of total Rac activation on an individual cell basis.

The labeling of PBD with Alexa, and mammalian expression vectors for expression of Rac-GFP is by any procedure, for example, as described in this application. This example provides protocols for production of pure PBD, protocols for generating cell images suitable for quantitative analysis of rac activation, and procedures and caveats for generating two types of data: images

showing the spatial distribution of Rac activation within cells, and curves showing the kinetics of Rac activation in single cells.

PBD Expression and Purification. PBD was expressed in the form of C-terminal 5 6His fusion from the prokaryotic expression vector pET23 (Novagen). It was determined experimentally that the highest levels of expression are observed when a vector containing plain T7 promoter (not T7*lac*) is used in combination with a BL21(DE3) strain (not the more stringent BL21(DE3)pLysS) of *E.coli*. This system allows for "leaky" protein expression (*Novagen*). While the 6His 10 tag can be cleaved from the purified protein with thrombin, it is not necessary, as the tag does not have any significant effect on probe functionality.

Competent BL21(DE3) cells (*Stratagene*) are transformed with pET23-PBD using standard procedures. See, e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press 1989) After 15 transformation, cells were plated on an LB plate containing carbenicillin. Cells do not degrade carbenicillin as quickly as ampicillin, so a higher percentage of cells retain the vector at the culture density appropriate for induction (*Novagen*). Five ml of LB media with 100 µg/ml carbenicillin were inoculated with a single colony of cells, and grown in the shaker at 37°C for 6-8 hours (until dense). Two 20 ml of this are then used to inoculate 50 ml of LBcarb. The rest of the culture is diluted 1:1 with glycerol and frozen for long term storage at -80°C. The 50 ml culture is incubated in the shaker overnight at 37°C. Next morning 1-2 L of LBcarb are inoculated with the overnight culture (15-20 mL culture/500 mL media), and grown in the shaker (37°C) to OD₆₀₀ = 0.8-0.9 (about 2-3 hours). 25 After that the cultures are briefly chilled on ice to 30-32°C, then put back in the shaking incubator turned down to 30-32°C. The protein is expressed at a lowered temperature to increase the portion of the correctly folded, soluble PBD. IPTG is added to a final concentration of 0.4-0.5 mM, and the cultures are allowed to grow for another 4-5 hours at 30-32°C (shaker). The cells are collected by 30 centrifugation (8,000 G, 4 min), and stored as a pellet at -20 °C until use. Approximately 2.5-3 g of cells is usually obtained from each liter of culture.

Purification of PBD-6His is performed essentially as described in the Talon affinity resin manual (*Clontech*). The cells (3-5 g) are thawed in 20-30 ml of the Lysis buffer (30 mM Tris HCl, pH 7.8, 250 mM NaCl, 10% glycerol, 5 mM MgCl₂, 2 mM β-ME, 1 mM PMSF), homogenized with a spatula and 5 sonicated (4 pulses, 10-15 sec each). T4 lysozyme and DNase are added in catalytic amounts (approximately 100 micrograms/ml lysozyme and 500 U DNase) to help the lysis, and the suspension is incubated on ice with periodic mixing for 30 min. The cells are then centrifuged at 12,500 rpm for 30 min, and the supernatant containing PBD is carefully transferred into a 50 mL Falcon 10 tube.

Talon resin (1.5-2 ml) (*Clontech*) is washed twice with 10 volumes of the lysis buffer in a 15 ml Falcon tube, centrifuging in the swinging bucket centrifuge at low speed in between to separate the resin. The cell lysate is added to the 1.5 ml of washed Talon resin in a 50 mL falcon tube, and inverted or 15 agitated gently (i.e. with an orbit shaker) for 20-30 mm at r.t. The resin is then separated by centrifugation in a swinging bucket centrifuge. The supernatant containing the unbound fraction is removed and saved for SDS-PAGE analysis. The resin is then transferred into a new 15 mL Falcon tube and washed twice (10-15 min each, r.t., orbit shaker) with 12 mL of the lysis buffer, without PMSF 20 and βME. The third wash is performed with lysis buffer + 10 mM imidazole (add 1 M stock in water, kept at -20°C). After the final separation, the resin is resuspended in 2-3 mL of lysis buffer with 10 mM imidazole, and pipetted into a column (0.5 cm in diameter). The resin is allowed to sediment by gravity flow until the fluid above the resin bed is almost gone, and then another 3-5 mL of 25 lysis buffer with 10 mM imidazole is added to wash the column. The elution is performed using Lysis buffer with 60 mM imidazole, and ca. 500 μL fractions are collected. PBD usually elutes in fractions 5-13 (total volume about 3-4 mL). An aliquot of each fraction is run on a 12% SDS-PAGE and the fractions containing the pure PBD are combined and dialyzed against 1 L of 25 mM NaP 30 buffer (pH 7.3). A dialysis bag (*SpectraPor 7*), or dialysis cassette (*Pierce*) with a molecular weight cutoff value of 3,500 kDa can be used.

- After 2-3 hours of dialysis, the bag is wiped with a KimWipe and buried in Aquacide powder (*Caibiochem*) for 15-45 min at 4°C, depending on the volume of the sample in the bag. This concentration process should be monitored carefully as complete drying may occur if the bag is left in the
- 5 Aquacide for too long. The powder is scraped gently from the bag every 10-15 min to facilitate water absorption. When the sample reaches 0.5-1.5 mL in volume (3-10-fold concentration), the Aquacide is cleaned from the bag, and the sample is carefully removed. The sample is briefly centrifuged (14,000 rpm, 2 min) to separate it from the precipitated material, and transferred into a new
- 10 dialysis bag or cassette. After the second dialysis step, the concentration of PBD is measured by taking a small aliquot (5-10 µL) and diluting into 50 mM TrisHCl (pH 7.5-8.0) or other appropriate buffer. The extinction coefficient of PBD at 280 nm is 8,250 (estimated from the primary sequence). On average, 1.5-2 mg of PBD is obtained per liter of cell culture.
- 15 Other methods of concentrating PBD were found to be less effective. For instance, centrifugal concentrators require prolonged centrifugations, and result in nonspecific adsorption of the small PBD protein to the membrane. It is essential to perform dialysis after concentration with Aquacide. This prevents the ionic strength of the resultant protein prep from becoming too high before
- 20 labeling. Low ionic strength conditions are preferable to avoid excessive precipitation of the protein during attachment of the hydrophobic dye.

Loading GFP-Rac and Alexa-PBD in cells. Cells were first transfected with GFP-Rac through nuclear microinjection. The EGFP variant (*Clontech*) was used that produced significantly brighter cells than wild type GFP (Heim et al., 6 Current Biology 178 (1996)). For microinjection of DNA and of PBD-Alexa, glass pipettes with 1.0 mm outer diameter and 0.50 mm inner diameter (*Sutter*) were pulled using a micropipette puller (*Sutter* Model P-87) to make microinjection needles with tips of approximately 0.5 µm diameter. Rac-GEP c-
30 DNA is injected into Swiss 3T3 fibroblasts at 200 ng/µL, using a constant needle pressure of approximately 100 hPa. DNA can be centrifuged prior to injection (20,000 G for 15 minutes) to prevent clogging the needle.

Cells expressing GFP-Rac were microinjected with Alexa-PBD using a microscope with optics and illumination capable of revealing the GFP fluorescence (detection sensitivity is typically improved by using higher NA objectives and brighter light sources, such as a 100W Hg arc lamp). Thus, only 5 GFP-expressing cells need to be injected. To reduce background fluorescence during injection and the following experiment, cells are placed in 1 mL of pre-warmed Dulbecco's Phosphate Buffer Solution (DPBS) containing 1000 mg/L D-glucose, 36 mg/L sodium pyruvate and supplemented with 0.2% BSA, 1% L-glutamine and 1 % Penicillin-Streptomycin. During injection and the following 10 experiment, cells are mounted in a Dvorak live cell chamber (*Nicholson*) preheated to 37°C, and maintained at 37°C by a heated stage (*20/20 Technology*). The microscope can be equipped with a motorized stage and shutter controls (*Ludi*) to monitor multiple stage positions in one experiment.

Cells that are barely expressing or expressing too much GFP-Rac were 15 ignored. The former produce FRET too weak for recording, and in the latter overexpressed Rac affects the biology of the cell. In general, we observed that cells expressing less than 300 intensity units (IU) do not display Rac-induced ruffling and altered morphology. The precise value of this cutoff will depend on the sensitivity of the imaging system, and should be determined by each lab for a 20 relevant biological behavior.

A 100 µM solution of Alexa-PBD was centrifuged at 20,000 G for 1 hour prior to injection, and then injected into the cytoplasm of cells expressing the GEP-Rac. Lowering the needle into the region just adjacent to the nucleus seems to produce the best combination of efficient injection and cell health. 25 After the injection, cells are placed back into the 37°C incubator for 5-10 minutes to recover. Alexa-PBD could potentially act as an inhibitor of Rac activity, so controls were carried out showing that, for our imaging system, up to 1000 IU of Alexa-PBD do not inhibit induction of ruffling.

30 Imaging Rac activation. Imaging experiments were performed using a *Photometrics KAF1400* cooled CCD camera, and *Inovision ISEE* software for image processing and microscope automation. Although filters are undergoing

further optimization, the best success to date has been achieved with the following filters, designed with sharp cutoffs specifically for this purpose by the Chroma corporation: GFP: HQ480/40, HQ 535/50, Q505LP FRET: D480/30, HQ610/75, 505DCLP Alexa: HQ545/30, HQ610/75, Q565LP.

5 The exact camera settings depend on the type of experiment being performed. When the total Rac activity within the cell is being determined, images are not generated, so spatial resolution can be sacrificed for increased sensitivity. Images are taken using 3x3 binning with exposure times of 0.1 sec, 0.1 sec, and 0.5 sec for GFP, Alexa and FRET respectively. When images are
10 required, i.e. to examine the changing spatial distribution of Rac activation, 1 x 1 binning is used with exposure times of 1 sec, 1 sec and up to 5 sec for GFP, Alexa and FRET respectively. These settings depend on the sensitivity of the imaging system used, and the desired trade off between sensitivity and spatial or temporal resolution. Settings should always be chosen not exceed the dynamic
15 range of the camera (Berland et al., in Sluder et al. (eds.), Video Microscopy at 33 (Academic Press 1998). Motion artifacts should also be considered during imaging experiments. For fast moving phenomena, features of the cell may move appreciably during the time between acquisition of the FRET and GFP images. This results in artifacts when the image is corrected for bleedthrough, as
20 described in more detail below. Such artifacts can be prevented by reducing the time between exposures, or by using two cameras simultaneously.

When total Rac activation is being determined, a picture of GFP-Rac and FRET is taken at each time point. Only one image of Alexa-PBD, usually at the initial time point, will also be required (for bleedthrough corrections as described
25 below). In contrast, when generating images (i.e. to determine the distribution of rac activation) an Alexa-PBD image is taken at each time point. The reasons for this are discussed in the 'Image Processing' section below. If the cells are to be treated with some type of stimulus, it is helpful to take a series of images prior to stimulation as controls for noise, bleaching, and other artifacts.

30

Image Processing. Image analysis is performed to follow the kinetics of total Rac activity within individual cells, displayed as curves of activation over time,

or to generate images that show the subcellular location of rac activation. The proper application of corrections is needed essential for quantitative imaging.

Common image processing operations can be used for this purpose (Berland et al., in Sluder et al., Video Microscopy at 33 (Academic Press 1998). Procedures

- 5 for such image processing operations will depend on the software package used. The correction factors should be rigorously applied when using the FLAIR system, as FRET signals will be low relative to other sources of fluorescence in the sample. The FRET signal must purposefully be kept low, as minimum quantities of fluorescent molecules should be used to prevent perturbation of cell
- 10 behavior. Hence, FLAIR procedures must be more carefully controlled than procedures generally used with fluorescent probes.

The following protocols assume use of a cooled CCD camera, which typically show low levels of noise, linear response to light intensity, and little variation in response from pixel to pixel. It is valuable to use cameras and

- 15 software with the greatest possible bit depth (allotment of computer memory to each pixel to maximize the number of possible intensity gradations). This is especially important for ratio operations, which are typically performed using 12 bit images or greater. Operations are described in the order in which they should be performed:

- 20 1. Registration. For corrections applied in later steps, it is important that each of the images taken using different excitation and emission wavelengths be registered so that the cells lie atop one another, with cell edges and internal features exactly coinciding. Different image processing software will accomplish this in different ways, but manual translocation is usually involved so that one
- 25 image lines up with a second, fixed image. This is best accomplished by zooming in on cells and adjusting brightness and contrast to clearly see cell edges and internal features. Since the GFP-Rac signal is generally the strongest, it was used as the reference image in these experiments. When the bleedthrough corrections described below are performed, errors in registration often become
- 30 apparent as 'shadow effects.'

2. Background subtraction. There are two methods commonly used for background subtraction. If the only intention is to follow the changing spatial

distribution of the FRET signal over time, and if the background (in the absence of cells) remains uniform across the field of view throughout the experiment, then it is sufficient to determine the background intensity in several regions of the image outside the cell. The average value of these intensities is then
5 subtracted from each pixel in the image. This method can also be sufficient for following qualitative changes in the subcellular location of activation, but it must be used with caution. Subtle variations in background intensity across the cell could be large relative to the changes observed in FRET, producing artifacts.
When quantifying the kinetics of total cell Rac activity over time (see following
10 sections), it is better to take an image of a region of the coverslip containing no cells or fluorescent debris, under the same conditions used for taking the real image. This background image is then subtracted from the real image prior to further analysis. A separate background image must be taken for each type of image (GFP, FRET or Alexa) and at each time point when successive images are
15 obtained

3. Masking. It is advisable to mask out regions surrounding the cells prior to further analysis. The edges of the cell are outlined, in most software either manually or by eliminating all sections of the images below a certain intensity value (interactive thresholding). The regions outside the cell are thus
20 identified and eliminated from further calculations. The precise approach for accomplishing this will depend upon the software employed. However, the mask is usually a binary image with all values within the cell = 1, and all outside = 0, and the real image is multiplied by the mask. The mask is best generated using the GEP image, which has the strongest signal and therefore the most clearly
25 defined edges. When determining the total intensity within the cell, analysis should be carried out on the same pixels within the FRET, GFP, and Alexa images. Therefore, the same mask is applied to each image after registration, assuring that exactly the same pixels are analyzed.

4. Bleedthrough correction. During FRET it is necessary to excite the
30 donor fluorophore while monitoring emission from the acceptor fluorophore. It is extremely difficult to design FRET filters that see only FRET emission and block all GFP emission, or block all light from Alexa excited directly rather than

- by FRET. To correct for “bleedthrough” of such light into the FRET image, the fluorescence filters must be characterized by taking images of cells containing GFP-Rac or AlexaPBD alone. For example, in bleedthrough correction for GFP, cells are imaged using both the GFP and FRET filter set. When observing the
- 5 GFP fluorescence through FRET filters, a fixed percentage of the GFP emission will be seen. The total fluorescence intensity is determined for both the GFP and FRET images from cells containing only GFP-Rac. A GFP bleedthrough factor is computed for each cell by dividing the intensity through the FRET filters by that through the GFP filters (the ‘bleedthrough factor’ for GFP: FRET
- 10 intensity/GFP intensity). This value is plotted against cell intensity for numerous cells, and a line is fit to this data to produce an accurate value of the bleedthrough factor. It is important to use background-subtracted images. The process is repeated for Alexa-PBD. When the actual experiment is performed, an Alexa-PBD, GFP-Rac, and FRET image are obtained. After background
- 15 subtraction of all three images, the Alexa-PBD and GFP-Rac images are multiplied by the appropriate bleedthrough factor and subtracted from the FRET image. This is an extremely important step that must be applied carefully to prevent artifacts that appear to be regions of high FRET, especially as the magnitude of the FRET signal approaches that of the bleedthrough. It is
- 20 important not to use GFP-Rac or Alexa-PBD images that exceed the dynamic range of the camera (‘overexposure’) as they will not fully eliminate bleedthrough. Motion artifacts can also produce errors derived from bleedthrough corrections.
- 25 Production of total activation curves (bleaching correction). To determine how overall Rac activity within a cell changes over time, the total fluorescence intensity within a GFP-Rac and FRET image is determined at each time point. The intensity of the FRET image is divided by that of the GFP-Rac image. This ratio better reflects total Rac activity than does FRET intensity alone. Division
- 30 by GFP-Rac normalizes out errors due to bleaching of GFP, effects of uneven illumination, and other factors affecting both the GFP and FRET signals. Because all FRET occurs through irradiation of GFP, bleaching of GFP will

decrease both GFP and FRET emissions to the same extent. Therefore, the FRET/GFP ratio will be a measure of Rac activity that is not affected by bleaching. It is critical that each image be properly background subtracted and corrected for bleedthrough. Bleedthrough corrections are somewhat simplified

5 when generating these curves, as they need not be carried out on actual images, but simply on the total intensity values derived from the images. The total intensity of the GFP-Rac image is multiplied by the GFP bleedthrough factor, and the resulting value is subtracted from the FRET intensity. An analogous operation is performed for Alexa-PBD bleedthrough. One need only obtain a

10 single Alexa-PBD image (usually at the beginning of the time series) and use this for bleedthrough correction of all images in the time series. If Alexa-PBD is not irradiated during the experiment, its bleaching will be negligible and this single image will reflect the actual Alexa-PBD level throughout the entire time series.

15 Any ratio calculations are best performed using floating point operations. Large errors can be generated by software that truncates noninteger values into integers to display data as images. Such problems can be overcome by multiplying the images by a large scalar value prior to division. This value should be as high as possible without causing any pixel to exceed the bit depth of

20 the image file. For example, a 12-bit image file can only hold values up to 4095 (2^{12}), so the constant selected must not cause the highest intensity value in the image to exceed 4095. When operations are performed on whole cell values rather than on images, as with production of the Rac activation curves described here, it is convenient to determine total cell intensities and then perform any

25 division operations in a floating point spreadsheet program, such as Microsoft Excel.

Examining changes in the localization of Rac activation. The sections above describe how to generate images showing the distribution of FRET within cells.

30 A sequence of such images can be compared to show how localizations vary with time. While simple examination can suffice to show distribution changes, the overall intensity of FRET, and hence the perceived activation level, could

become successively lower over time due to bleaching of the GFP. We have found that bleaching is not a serious issue for at least 20 images under the exposure conditions described here. The total GFP intensity at the beginning and end of the experiment should be examined to gauge bleaching effects. One 5 can correct for bleaching by dividing each pixel in the image by the total intensity of GFP determined at the same time point.

Many aspects of cytoskeletal control and signaling crosstalk depend upon the localization of Rho family GTPase activation, and may depend as well on the level and duration of activation. Signaling control by the precise dynamics of 10 GTPase activation has been suggested by indirect experiments, but has been very difficult to quantify or study using previous methods. The FLAIR system reported here can reveal Rac activation dynamics *in vivo* and can accurately report the changing activation levels within a cell.

15 **Example 7. Cellular Localization of Rac GTPase Activation Using p21 Activated Kinase (PAK) as a Protein Biosensor**

Prevailing evidence suggests that signaling proteins must be tightly regulated both spatially and temporally in order to generate specific and 20 localized downstream effects. For Rac1 and other small GTPases, binding to GTP is a critical regulatory event that leads to interaction with downstream targets and regulates subcellular localization. Here we use FLAIR (fluorescence activation indicator for Rho proteins), to quantify the spatio-temporal dynamics of Rac1 GTP interaction in living cells. FLAIR reveals precise spatial control of 25 growth factor-induced Rac activation in membrane ruffles, and in a gradient of activation at the leading edge of motile cells.

Rac is a member of the Ras superfamily of small GTPase proteins (A. Hall, Science 279, 509-14 (1998). It plays a critical role in diverse signaling pathways, including control of cell morphology, actin dynamics, transcriptional 30 activation, apoptosis signaling, and other more specialized functions (L. Kjoller, A. Hall, Exp. Cell Res. 253 166-79 (1999). The broad range of events controlled by this GTPase requires subtle regulation of interactions with multiple downstream targets. Accumulating evidence suggests that the effects of Rac are

in part controlled by regulating the subcellular localization of its activation through GTP binding. For example, Rac is known to induce localized actin rearrangements to generate polarized morphological changes (C.D. Nobes, A. Hall, *J. Cell Biol.* 144(6), 1235-1244 (1999). GTP exchange factors (GEFs) 5 which regulate nucleotide exchange on Rho GTPases contain a variety of localization domains and may modulate downstream signaling from Rac (Zhou et al, *J. Biol. Chem.* 273(27), 16782-16786 (1998).

Although control of Rac1 signaling through localized activation is clearly important, it is difficult to study in an intact cell. Here we develop and apply a 10 new method based on fluorescence resonance energy transfer (FRET) which can quantify the timing and location of Rac activation through GTP binding. This method is applied to examine the hypothesis that specific and localized Rac1 activation occurs during the course of extracellular signal-induced cytoskeletal dynamics.

15 The design of the Rac nucleotide state biosensor is shown schematically in Figure 5. A fluorescently labeled protein biosensor is introduced into the cell together with a biologically active GFP-fused Rac (C. Subauste et al, *J. Biol. Chem.* 275(13), 9725-9733 (2000). This protein biosensor is labeled with an acceptor dye capable of undergoing FRET with GFP. Since the biosensor is 20 derived from a specific GTP-Rac target protein, it binds to GFP-Rac 1 only when the Rac1 is in its activated, GTP-bound form, and produces a localized FRET signal revealing the level and location of Rac activation. When cells expressing GFP-Rac were injected with the biosensor, we were able to simultaneously map the changing location of GFP-Rac1 and the subpopulation of GFP-Rac1 25 molecules in the activated, GTP-bound state. FRET is proportional to the amount of GTP binding, enabling quantitation of changing activation levels. It is also very specific, as only biosensor binding to the GFP-tagged protein will generate FRET. This approach has the potential to examine many protein states, including posttranslational modifications, conformation, and ligand binding.

30 The biosensor was made by fluorescently labeling a domain from p21 activated kinase I (PAK1) known to bind selectively to GTP-Rac . PBD (fragment of human PAK1, residues 65-150, with a single cysteine added in the

penultimate N-terminal position) was expressed as a C-terminal 6His fusion protein from the pET23 vector (Novagen) and purified from *E. coli* strain BL21DE3 using Talon metal affinity resin (Clontech) as instructed by the manufacturer. All GFP constructs were prepared using the EGFP mutant (T.

- 5 Joneson, *Mol. Cell. Biol.* 19(9) 5892-901(1999); T.T. Yang et al., *Gene* 173, 19-23 (1996), kindly provided by Dr. Roger Tsien of the University of California, San Diego. GFP-Rac 1 fusion and wild type Rac1 for *in vitro* studies were also expressed as 6His constructs and purified using a similar procedure. Purified protein was dialyzed against 50 mM sodium phosphate (pH 7.8), and labeled
10 with 7 equivalents Alexa 546 maleimide (Molecular Probes) at 25 degrees for 2 hours. The conjugate was purified from unincorporated dye by G25 size exclusion chromatography followed by dialysis. The dye:protein ratio was between 0.8 and 1.3, as determined from absorbance of the conjugate at 558 nm (Alexa 546 extinction coefficient 104,000 M⁻¹cm⁻¹) and 280nm (PBD, extinction
15 coefficient 8,250 M⁻¹cm⁻¹ plus Alexa absorbance, determined as 12 % of the absorbance at 546). Protein concentration was also independently determined using a Coomassie Plus protein assay (Pierce) and SDS-PAGE calibration with known concentrations of bovine serum albumin.

The p21 binding domain (PBD, aa 65-150) has been used successfully to
20 precipitate GTP-Rac 1 from cell lysates (V. Benard, B.P. Bohl, G.M. Bokoch, J. Biol. Chem. 274, 13198-204(1999). In order to produce efficient FRET, the optimum site for attachment of an acceptor dye was determined by analyzing FRET between purified GFP-Racl and PBD labeled with various dyes in different positions. PBD contained no native cysteines, so the site of labeling
25 could be controlled through introduction of a single cysteine, followed by labeling with cysteine-selective iodoacetamide dyes. After considerable optimization, the best candidate was found to be a PBD with cysteine appended to the N-terminus, labeled with commercially available Alexa 546 dye. The distance between the Alexa dye at the N-terminus of PBD and the fluorophore in
30 GFP was calculated to be 52 Å based on the efficiency of FRET and assuming random rotation of the fluorophores ($R_0 = 51$, $n = 1/4$, $k^2 = 2/3$) (T. Nomanbhoy et al., *Biochemistry* 35, 4620-4628 (1996). We coined the name FLAIR

(fluorescent activation indicator for Rho proteins) for this new live cell imaging technique.

FRET between Alexa-PBD and GFP-Rac1 was efficient and dependent on GTP-Rac 1 binding. Using fluorescence excitation wavelengths that selectively excite GFP (480 nm), fluorescence emission was monitored while maintaining a fixed concentration of GFP-Rac 1 and varying AlexaPBD concentrations (see Figure 7a). Purified GFP-Rac1 (200 nM) was bound to varying concentrations of GTP γ S or GDP at low magnesium by 30 min incubation at 30°C as described previously (U.G. Knaus et al., *J. Biol. Chem.* 10 267, 23575-23582 (1992)), using the following nucleotide equilibration buffer: 50 mM Tris HCl (pH 7.6), 50 mM NaCl, 5 mM MgCl₂, 10 mM EDTA, and 1 mM DTT. Equal volumes of Alexa-PBD in the same buffer were added to the GFP-Rac1 solution and fluorescence emission spectra (500-600 nm) were acquired at room temperature and 480 nm excitation. Spectra were corrected for dilution upon Alexa-PBD addition. Alexa-PBD concentrations were either varied as shown, or maintained at 1 micromolar when saturating Alexa-PBD was required. The spectra shown were corrected for direct excitation of the Alexa fluorophore by acquiring spectra of Alexa-PBD alone at equivalent concentrations, and subtracting these from spectra shown in Figure 7. K_d were determined by fitting to the equation: Y = A*X/(K_d + X). In the Figure 7, panel A inset, only points actually used in the curve fitting are shown. Higher, saturating concentrations of AlexaPBD were not used because errors from subtraction of direct Alexa excitation became larger. Binding of Alexa-PBD to GFP-Rac resulted in a change in fluorescence intensity of both donor (GFP) and acceptor (Alexa) emission. When the GFP and Alexa fluorophores were brought into close proximity, FRET caused the Alexa (acceptor) emission to increase while the GFP (donor) emission decreased. No change in emission was observed when using either unlabeled PBD or Rac1 not fused to GFP, indicating that fluorescence changes were in fact due to energy transfer (data not shown).

Because FRET alters donor and acceptor intensities in opposite directions, the ratio of emission at these two wavelengths is a sensitive measure of the PBD-Rac1 interaction. The corrected Alexa/GFP emission ratio underwent a 4-fold

change upon saturation of Rac 1 with GTP (Figure 7b). The change in emission ratio versus PBD concentration was fit to the Michaelis equation to derive an apparent K_d for PBD-Rac1 binding of $1.1 \pm 0.3 \mu\text{M}$ (Figure 7a inset). This is slightly higher than the values that have been determined for various unlabeled

5 PAK1 fragments (E. Manser, T. Leung, H. Salihuddin, Z.-S. Zhao, L. Lim, Nature 67, 40-46 (1994); D.A. Leonard et al, Biochemistry 36, 1173-1180 (1997); G. Thompson, D. Owen, P. A. Chalk, P. N. Lowe, Biochemistry 37, 7885-7891 (1998). This may indicate that the presence of the Alexa may somewhat weaken the binding interaction. Our studies in live cells, described

10 below, demonstrated that this affinity was sufficient for detection of Rac activation in vivo, and reversible binding is in fact desirable as it permits the biosensor to rapidly respond to changes in the location or level of Rac activation. Importantly, Alexa-PBD was shown to minimally perturb Rac-GTP binding interactions. The apparent $\text{GTP}\gamma\text{S}$ dissociation constant was determined at

15 saturating Alexa-PBD by fitting the experimental data to the Michaelis equation (Figure 7b). The derived value of $47 \pm 9 \text{ nM}$ is consistent with the previously reported value of 50 nM (L. Menard, E. Tomhave, P.J. Casey, R.J. Uhing, R. Snyderman, J.R. Didsbury, Eur. J. Biochem. 206, 537-546 (1992)). This validated application of FLAIR as an indicator of biologically relevant Rac-

20 nucleotide binding.

Much is known about the localized dynamics of actin, but it has been difficult to explore how signals can generate precisely localized actin behavior. Rac has been shown to participate in signaling cascades leading to localized polymerization in several systems, but it is unclear whether Rac activity is

25 constrained to specific subcellular regions, or whether global activation of Rac leads to more localized activation of downstream molecules. The precise spatial correlation of localized signaling and downstream actin behavior remains completely unknown. Here, we used FLAIR to study the localization and activation of Rac in two cell systems where actin behaviors are clearly

30 constrained to specific subcellular regions, the induction of ruffling by growth factor stimulation of quiescent cells, and polarized cell motility induced by wounding.

We examined stimulation of quiescent Swiss 3T3 fibroblasts by serum or platelet derived growth factor (PDGF) known to initiate membrane ruffling and transcription through activation of Rac 1 (Ridley A.J., Paterson H.F., Johnston C.L., Diekmann D., Hall A., Cell 70,401-10 (1992); P.T. Hawkins, Curr. Biol. 5(4), 393-403 (1995)). For serum stimulation experiments, Swiss 3T3 fibroblasts (ATCC, passage 15-27) were plated on glass cover slips and then maintained in Dulbeccos modified Eagle's medium (DMEM) with 10% fetal calf serum (FCS), 1% L-glutamine and 1% penicillin-streptomycin for at least 24 hours. Media was then replaced with media containing only 0.5% FCS, and 10 cells were maintained for 42 hours. Cells were transfected by microinjecting 200 micrograms/ml GFP-Rac1 c-DNA into cell nuclei 2-8 hours prior to the experiment. The EGFP mutant was used in all experiments, cloned and expressed as previously described (C. Subauste et al, J. Biol. Chem. 275(13), 9725-9733 (2000)). Cells expressing the GFP-Rac were then microinjected with 15 100 micromolar Alexa-PBD, mounted in a heated chamber on a Zeiss axiovert 100TV microscope and maintained in Dulbecco's phosphate buffered saline (DPBS) (GIBCO) to reduce background fluorescence. Cells were then stimulated by replacing the media with DPBS containing 10% FCS or 50 ng/mL PDGF. Images were obtained every 30 seconds using a Photometrics PXL cooled CCD 20 camera with 1x1 or 3x3 binning, and a Zeiss 40x 1.3 NA oil immersion objective. Fluorescence filters from Chroma were as follows: GFP: HQ480/40, HQ535/50, Q505LP; FRET: D480/30, HQ610/75, 505LP; Alexa:HQ 545/30, HQ 610/75, Q565LP. Cells were illuminated using a 100W Hg arc lamp. Exposure times for 3x3 binning were: GFP- 0.1 seconds, Alexa-PBD - 0.1 seconds, FRET -0.5 25 seconds. For 1x1 binning, GFP - 1 second, Alexa-PBD - 1 second, FRET - 5 seconds.

Levels of exogenous proteins had to be limited to concentrations that would not perturb cell behavior. We determined the intracellular levels of Alexa-PBD and GEP-Rac 1 that altered normal serum-induced ruffle formation 30 (Figure 8), and kept exogenous protein below these levels throughout our studies.

Image triplets of GFP, FRET, and Alexa fluorescence were taken at each successive time point before and after stimulation. Thus, as shown in Figure 9, both the changing localizations of GFP-Rac, and the level and location of Rac activation could be monitored. Images were first background subtracted and

5 carefully registered to ensure accurate pixel alignment. The GFP-Rac image was then thresholded, changing the intensities of all pixels outside of the cell to zero. Thresholding was based on the GFP image since it had the largest signal to noise ratio, providing the clearest distinction between the cell and background. The thresholded GFP-Rac image was used to generate a binary image with all values

10 within the cell = 1 and all outside = 0. The FRET and Alexa-PBD images were multiplied by the binary image, assuring that exactly the same pixels were analyzed in all three images. Emission appearing in the FRET image from direct excitation of Alexa and GFP was removed by subtracting a fraction of the GFP-Rac and Alexa-PBD images from the FRET image. This fraction depended on

15 the filter set and exposure conditions used. It was determined, as described in detail elsewhere (C.E. Chamberlain, V. Kraynov, K.M. Hahn, *Methods Enzymology* (In Press 2000)), by taking images of cells containing only GFP-Rac or Alexa-PBD alone, and quantifying the relative intensity of emission in the FRET channel and that in the GFP or Alexa-PBD channel. A broad range of

20 intensities was examined and a line was fit to these for accurate determinations. These corrections had to be applied carefully when studying rapidly moving objects such as ruffles. If the ruffle moved between acquisition of the FRET, GFP, or Alexa images, the subtractive correction process would remove light from the FRET image in the wrong place, generating artefactual FRET

25 localizations. Data from moving features was used only when careful inspection showed the feature to be coincident in the Alexa, GFP, and FRET images, and controls were performed with images taken in different orders. A low pass filter kernel was applied to the corrected FRET image to remove high frequency noise (K. Castleman, *Digital Image Processing* (Prentice Hall, New Jersey, 1996),

30 pp.207-209). Image processing and microscope automation were performed using Inovision ISEE software. Images were contrast stretched and formatted for display using Adobe Photoshop software.

We tested Rac and PBD fused to GFP mutants that undergo FRET (ECFP and EYFP) (S. Dharmawardhane, D. Brownson, M. Lennartz, G.M. Bokoch, *J Leukoc. Biol.* 66(3), p. 521- 527 (1999); B. A. Pollok and R. Heim, *Trends Cell. Biol.* 9(2), 57-60 (1999)). Unfortunately, their spectral overlap was 5 more problematic than that of Alexa and GFP, making the corrections described here more difficult. In addition, the GFP mutants showed roughly 25% the FRET of FLAIR. We decided to use FLAIR for these reasons, but the ability to monitor Rac activity simply through protein expression may justify using GFP mutants in some applications.

10 Simple GFP-Rac fluorescence revealed pools of Rac1 at the nucleus, in the juxtanuclear region, and in small foci throughout the cell prior to stimulation. Confocal and deconvolution imaging showed the nuclear Rac to be concentrated at the nuclear envelope, and expression and immunostaining of HA-tagged Rac1 indicated that these localizations were not an artifact of GFP tagging. Addition 15 of PDGF or serum led to formation of moving ruffles throughout the cell periphery within 2 minutes. These contained GFP-Rac, shown by phalloidin staining of fixed specimens to colocalize with actin (data not shown). The FRET images showed a stark contrast between the level of Rac activation in the ruffles and the nucleus. No FRET was seen at the nucleus despite the high 20 concentration of Rac1 there, while the moving ruffles showed the highest FRET, clearly restricted to the ruffle and discernable from the rest of the cell. The Rac activation remained tightly correlated with the position of the ruffle even as it moved throughout the cell.

As a negative control, FRET was imaged in cells expressing a mutant of 25 GFP-Rho, a close relative of Rac that should not bind PBD. This GFP-Rho Q63L mutant, which generates high levels of GTP-bound protein, produced clear Rho localizations but no corresponding FRET signals (data not shown).

These studies provide the first direct evidence that Rac1 activation is 30 restricted to the site of actin polymerization, independent of the overall distribution of the protein. Rac activation was tightly correlated with moving ruffles, indicating that structures specifically associated with the ruffle were either binding and concentrating activated Rac or that growth-factor induced Rac

activation was specifically localized to ruffles. The function of the Rac found at the nuclear envelope remains uncertain. It may be activated for regulation of transcription at times later than those tested here, or may be activated for an unknown role by other stimuli. When activation was concentrated in a small area

5 such as a ruffle, spatially resolved FRET could detect significant activation changes too small to appreciably alter the overall levels of cellular Rac activity.

Our data showed that FRET provided much greater sensitivity and selectivity than following Rac activation simply by imaging Alexa-PBD localization (Figure 9, panels C and D). FRET produces much lower backgrounds and

10 provides complete selectivity even when the biosensor can bind to multiple proteins (i.e. Alexa-PBD also binds to cdc42). Unlike simple localization, FRET can provide quantitative measures of activation levels. It is important to note that the Alexa-PBD could have been sterically hindered from reaching Rac1 in some locations, so that not all activated Rac1 may have been detected. Nonetheless, a

15 FRET signal in a given location does reveal that Rac activation is occurring there.

Rac has been shown to be essential for the directed movement of cells during chemotaxis, and for extension of the front end of cells during motility (C.Y. Chung et al, Proc. Natl. Acad. Sci. U.S.A. 97(10), 5225-5230 (2000). We

20 used FLAIR to ask if Rac activation in polarized, motile cells occurred in particular subcellular localizations to regulate localized actin behaviors. A 'wound' was scraped in a monolayer of confluent Swiss 3T3 fibroblasts, causing cells to become polarized and move into the open space. For wound healing experiments, Swiss 3T3 fibroblasts were induced to undergo polarized

25 movement as previously described (R. DeBiasio, G.R. Bright, L.A. Ernst, A.S. Waggoner, D.L. Taylor, *J. Cell. Biol.* 105, 1613-22 (1987). The cells were cultured in Dulbecco's modified Eagle's medium (GIBCO) supplemented with 10% fetal calf serum at 37°C. Cells were trypsinized and then plated on glass coverslips. They were grown to a confluent monolayer and maintained for an

30 additional 3-4 days. Cells were then wounded by creating a straight laceration with a sterile razor blade. Cells along the edge of the wound were microinjected with 200 micrograms/ml GFP-Rac c-DNA. Six hours after the wound was

formed, cells expressing the GFP-Rac were microinjected with 100 micromolar Alexa-PBD and allowed approximately 10 minutes for recovery. Media was then replaced with DPBS containing 10% FCS to reduce background fluorescence. Images were obtained as described above, using exposure times of 5 1 second for GFP, 1 second for Alexa-PBD, and 5 seconds for FRET. FLAIR revealed highest Rac activation in the juxtanuclear area, and a gradient of Rac activity highest near the leading edge and tapering off towards the nucleus (Figure 10).

Quantitative analysis clearly indicated that the gradient was correlated 10 with the direction of movement. In individual cells, total Rac activity was measured in two subcellular locations: where activity was highest at the front of the cell, and lowest at the very rear of the cell. These values, measured in squares three microns on a side, were used to calculate the gradient (Percent gradient = $100 \times [\text{front} - \text{back}] / \text{back}$). Of the 16 cells examined, twelve had 15 higher Rac activity at the leading edge, while four showed a slight negative gradient. For cells with high Rac activity at the leading edge, the gradient was $128 \pm 51\%$, while for cells showing the reverse gradient, it was only $9 \pm 4\%$ (mean \pm standard error). The gradient was much broader than the narrow area at the leading edge where actin polymerization occurs (Y.L. Wang et al, *J. Cell Biol.* 101, 597-602 (1985; J.A. Theriot and T.J. Mitchison, *Nature* 352, 126-131 (1991). Other activities required for motility, such as depolymerization of fiber networks to recycle monomers and delivery of molecules to the leading edge (O.D. Weiner et al., *Nat. Cell. Biol.* 1, 75-81 (1999) occur throughout the region where Rac is activated. Perhaps Rac is acting over a broader cell area to activate 20 multiple downstream effectors, each producing different effects in more restricted locations. Other studies have shown tight localization of molecules downstream of Rac, at the leading edge or in regions immediately behind it to regulate a variety of functions associated with motility (F. Michiels et al., *Nature* 375, 338-340 (1995).

25 30 The prevalence of Rac activation around the nucleus was quantified by scoring 16 cells. All cells showed both juxtanuclear and nuclear GFP fluorescence. Of these, fourteen showed juxtanuclear FRET, and none showed

nuclear FRET. It was noteworthy that small areas of the nucleus sometimes showed a FRET signal, but these could be due either to cytoplasmic Rac associated with the nuclear envelope or to juxtanuclear localizations lying over the nucleus. The localization of activation within the juxtanuclear Rac often did

- 5 not parallel Rac distribution, with "hot spots" of FRET within areas of lower Rac concentration. The meaning of the juxtanuclear Rac localizations is unclear, but their morphology and distribution suggests activation within the ER, golgi, or vesicle populations, perhaps consistent with recent reports suggesting an important role for Rac in ER to golgi transport (M.P. Quinlan, *Cell Growth Diff.* 10(12), 839-854 (1999), and in pinocytic vesicle cycling (Ridley A.J., Paterson H.F., Johnston C.L., Diekmann D., Hall A., *Cell* 70,401-10 (1992).

In summary, we have described a novel approach to quantify the spatial distribution and rapidly changing levels of Rac signaling in living cells. FLAIR provided the first direct demonstration that Rac activity is spatially regulated to 15 generate specific actin behaviors in different subcellular regions. Activated Rac was tightly coupled to small membrane ruffles even as the ruffles moved throughout the cytoplasm, yet was broadly distributed as a gradient at the leading edge of motile cells. These very different activation patterns suggest that the cell will utilize different distributions of activated Rac depending on the number and 20 localization of downstream targets.

Ultimately, FLAIR can reveal how different stimuli interact to affect Rac through the complex circuitry of an intact cell. We have focused here on spatial control of signaling. The ability of the biosensor to quantify the level and kinetics of activation should also prove very useful, as accumulating evidence 25 indicates that Rac and related proteins do not function as simple binary switches. Different levels and kinetics of activation produce profoundly different results (T. Joneson, *Mol. Cell. Biol.* 19(9) 5892-901(1999). Using FLAIR together with other biosensors of different wavelengths it should be possible to examine the balance between Rac, Ras, and other protein activation levels undergoing rapid 30 changes in real time. With increasing access to FRET imaging equipment, the technique we describe provides a relatively straightforward way to greatly extend the utility of readily accessible GFP fusion proteins.

Example 8: Activation biosensors for cdc42 and Erk2.

Like Rac, cdc42 is a member of the Rho family of small GTPases involved in signal transduction in eukaryotic cells. Cdc42 becomes "activated" 5 by releasing GDP and binding GTP. Such GTPases interact with a host of downstream effectors, ultimately resulting in one or the other cellular response via a variety of phosphorylation cascades. In these experiments, a unique synthetic fluorophore with environment-sensitive fluorescence properties was linked to the cdc42 binding domain ("CBD") of the Wiscott-Aldrich Syndrome 10 Protein (WASP) to generate a CBD biosensor. Upon binding to the cdc42, this fluorescent CBD biosensor is able to increase its fluorescence intensity by up to 3.5-fold, providing a convenient measure of endogenous cdc42 activation in living cells or for in vitro applications (concept outlined in figure 6b).

Mitogen-activated protein kinase (MAP kinase) was first identified as a 15 protein phosphorylase which is activated when a growth factor is added to cultured cells (Proc. Natl. Acad. Sci. USA, 84, 1502-1506, 1987). However, subsequent research revealed that this enzyme is involved in various vital phenomena such as neuronal differentiation (J. Biol. Chem., 265, 4730-4735, 1990), activation of immune cells (J. Immunol., 144, 2683-2689, 1990) and 20 secretions (J. Cell Biol., 110, 731-742, 1990). MAP kinase is also called Extracellular Signal-Regulated Kinase, or ERK. Cloning of the gene for human ERK gene showed that it consists of several molecular species with high homology, but the main species are two, namely ERK1 and ERK2. These two proteins are highly homologous (84.7%) (FEBS LETT., 304, 170-178, 1992). 25 Mitogen-Activated Protein Kinase Kinase (called MEK) can interact with and stimulate ERK2.

Experimental Protocols:

Production of recombinant proteins. DNA encoding the Cdc42-binding 30 fragment of human WASP containing the CRIB motif and surrounding amino acids (WASP amino acids 201 to 321) was amplified by PCR from ATCC clone

99534. This peptide fragment has the following amino acid sequence (SEQ ID NO:13).

5 DIQNPDTSSRYRGLPAPGPSPADKKRSGKKKISKADIGAPSGFKHVSHV
GWDPQNGFDVNLDPLRSLSRAGISEAQLTDAETSKLIYDFIEDQGGL
EAVRQEMRRQEPLPPPPPS

The full sequence of the WASP protein is as follows (SEQ ID NO:14):

10 MSGGPMGGRP GGRGAPAVQQ NIPSTLLQDH ENQRLFEMLG RKCLTLATAV
VQLYLALPPG AEHWTKEHCG AVCFVKDNPQ KSYFIRLYGL QAGRLLWEQE
LYSQLVYSTP TPFFHTFAGD DCQAGLNFAD EDEAQAFRAL VQEKIQKRNQ
RQSGDRRQLP PPPTPANEER RGGLPPLPLH PGGDQGGPPV GPLSLGLATV
DIQNPDTSS RYRGLPAPGP SPADKKRSGK KKISKADIGA PSGFKHVSHV
15 GWDPQNGFDV NNLDPLRSLS FSRAVISEAQ LTDAAETSKLI YDFIEDQGGL
EAVRQEMRRQ EPLPPPPPS RGGNQLPRPP IVGGGNKGRSG PLPPVPLGIA
PPPPTPRGPP PPGRGGPPPP PPPATGRSGP LPPPPPGAGG PPMPPPPPPP
PPPPSSGNGP APPPLPPALV PAGGLAPGGG RGALLDQIRQ GIQLNKTPGA
PESSALQPPP QSSEGLVGAL MHVMQKRSRA IHSSDEGEDQ AGDEDEDDEW
20 DD

The DNA fragment encoding SEQ ID NO:13 was subcloned into pET23a (*Novagen*) as a C-terminal 6His fusion. Site-specific cysteine mutants were constructed by QuikChange (*Stratagene*) mutagenesis using synthetic oligos and
25 the presence of mutations was confirmed by DNA sequencing. Resultant constructs were transformed into BL21DE3 strain of *E.coli* (*Novagen*), and the proteins were produced by expression at 30°C for 5 hours in 1 L Leuria-Bertani media (*Sigma*) in the presence of 100 µg/ml of carbenicillin. Expression was induced with 0.5 mM IPTG at OD₆₀₀=0.8-1. Cells were collected by
30 centrifugation and stored at -20°C until use.

Cell pellet was resuspended in cold lysis buffer (25 mM Tris-HCl, pH 7.9, 150 mM NaCl, 5 mM MgCl₂, 5% glycerol, 1 mM PMSF, 2 mM β-mercaptoethanol), and briefly sonicated on ice. Lysozyme and DNase were added to the suspension to a final concentration of 0.1 mg/ml and 100 U/ml,
35 respectively, and solution was incubated with occasional stirring at 4°C for 30 min. Lysate was centrifuged (12,000 g, 30 min), and the clarified supernatant was incubated with 1 ml of Talon resin (*Clontech*) at 25°C for 30 min. The resin

containing bound CBD was separated from the lysate by brief low-speed centrifugation and washed twice with 15 ml of lysis buffer. Finally, the resin was washed with the lysis buffer, supplemented with 10 mM imidazole, and poured into a column. Elution was performed with 5 ml of the lysis buffer, containing 5 60 mM imidazole. Fractions containing bulk of CBD (as evidenced by SDS gel) were combined and dialyzed against 1 L of dialysis buffer (25 mM Na₂HPO₄ (pH 7.5), 10 mM NaCl) for 5 hours at 4°C. Solution was then concentrated using Aquacide powder to a final protein concentration of 2 to 10 mg/ml, and dialyzed once again against dialysis buffer. Final preparation was flash-frozen in 100 µL 10 aliquots and stored at -80°C. Generally, 1 to 3 mg of CBD was obtained from 1 L of cell culture. Recombinant 6His-tagged cdc42, RhoA, Rac1, ERK2 and MEK were produced by analogous procedures. The enzymes were determined to be >90% active by the GTP binding assay (Knaus, U.G., Heyworth, P.G., Kinsella, B.T., Curnutte, J.T., and Bokoch, G.M. (1992) *J. Biol. Chem.* 267, 15 23575-23582).

Conjugation of CBD with fluorescent dyes - Dialyzed CBD samples (100-150 µM) were gently inverted with 6 to 7-fold molar excess of the reactive dye at 25°C for 3 to 4 hours. The reaction was stopped by addition of 10 mM dithiothreitol (DTT), and the mixture was incubated for 15 min. Unreacted dye 20 was separated from the labeled protein using G25-Sepharose (Pharmacia) gel filtration column equilibrated and developed with 25 mM Na₂HPO₄ (pH 7.5). Purity of the eluting fractions was analyzed by running an aliquot on an SDS gel and visualizing the fluorescence. Only the fractions containing minimal amounts of free dye were used in the subsequent experiments. Dye-to-protein ratio was 25 determined by measuring CBD concentration ($\epsilon^{280}=8,250\text{ M}^{-1}$), and A4C concentration at 617 nm ($\epsilon=70,000\text{ M}^{-1}$ in dimethylsulfoxide) or Alexa546 at 554 nm ($\epsilon=104,000\text{ M}^{-1}$ in 50 mM potassium phosphate, pH 7.0). Concentrations of CBD were independently confirmed by Coomassie Plus assay (Pierce) calibrated with bovine serum albumin as a standard. Dye-to-protein 30 ratios thus obtained varied between 0.8 and 1.2, 1.7 and 2.1 for the single-dye and dual-dye conjugates, respectively. Aliquots of the labeled CBD (15 to 50 µM) were stored at -80°C. No significant loss of binding ability was observed

after 6 months of storage. In this example, CBD-conjugates were made with the dye shown in Figure 11, and are referred to as mero-CBD (for merocyanines dye conjugated to CBD).

5 Dye Attachment Single Site on ERK2 - The ERK2-dye adduct was produced using conditions which should label only the single solvent-exposed cysteine on ERK. This exposed cysteine was located away from known protein interaction sites. A maleimide reactive group on the dye was attached to this cysteine residue at pH 7.5. The attachment of label and the activation of ERK2
10 was examined under different conditions (Figure 23). ERK2 was activated with MEK for 0-60 min, as indicated in Figure 23. When Mg and ATP were added, ERK2 was phosphorylated; no such phosphorylation was observed when no Mg, ATP or MEK was present (Figure 23).

15 Fluorescence increase as a measure of ERK2-MEK interaction - The functionality and specificity of ERK2 biosensor was characterized by measuring fluorescence in the presence and absence of saturating amounts of ATP (Figures 24 and 25). An approximate 1.6-fold increase in emission was observed, when the Erk2 was phosphorylated (Figure 24). No effect on Erk2 fluorescence was
20 observed when no MEK was present (Figure 25). This result demonstrated the suitability of the present methods for detecting activation of ERK2 in live cells.

25 Analyses of the solvatochromic activation indicator - A solution of mero-CBD (300 nM) in assay buffer was mixed 1:1 (v/v) with solutions of cdc42 (concentration as indicated), pre-equilibrated with 10 mM GDP or GTP γ S as described in Knaus, U.G., Heyworth, P.G., Kinsella, B.T., Curnutte, J.T., and Bokoch, G.M. (1992) *J. Biol. Chem.* 267, 23575-23582. Emission wavelength of 630 nm and excitation wavelength of 600 nm were used to acquire excitation and emission spectra, respectively. For nucleotide dependence, cdc42 (500 nM)
30 was pre-incubated with varying concentrations of GTP γ S (1 to 500 nM). Rac1 and RhoA GTPases were pre-equilibrated with GTP γ S in the same manner.

Assay of cdc42 activity in stimulated neutrophil cell lysates. - Freshly prepared neutrophils (2.5×10^7 cells/sample) were incubated in a Krebs-Ringer HEPES buffer supplemented with 5.5 mM glucose (KHRG), in the presence of 1 mM CaCl₂ and 1 μM of fMet-Leu-Phe tripeptide (fMLP) for various periods of time at 37°C (see Benard, V., Bohl, B.P., and Bokoch, G.M. (1999) *J. Biol. Chem.* 274, 13198-13204). Stimulation was stopped by adding equal volume of 2X lysis buffer, supplemented with 2 % Nonidet P-40, 2 μg/ml aprotinin. After a brief vortexing the lysates were clarified by centrifugation and immediately analyzed for cdc42 activity with mero-CBD as described above. In control samples, the lysates were pre-equilibrated with either GDP or GTPγS at 30°C for 20 to 30 min. This was a convenient method to determine the extent of cdc-42 activation in cell lysates, as shown in Figure 16.

Using the structural information available (Abdul-Manan et al. (1999) *Nature* 399, 379-383; Kim et al. (2000) *Nature* 404, 151-158), we had selected three positions for placing the fluorescent dye within the CBD peptide (I233, D264 and F271, by WASP numbering), that could experience a considerable change in solvent polarity as a result of binding to cdc42 (see Figure 14). F271 and D264 are located in the loop contacting the effector “switch” domain of the GTPase, whereas I233 appears to interact with the hydrophobic pocket formed by residues at the N-terminus of cdc42 (see Figure 14). The three CBD residues were mutated to cysteines, easily amenable to covalent modification with the thiol-reactive derivatives of the solvatochromic dyes. Recombinant mutant CBD proteins were overexpressed in bacteria, purified and site-specifically modified with several of solvatochromic dyes. Only the conjugates with the dye shown in Figure 11 are described here.

Among the three mutants tested, the mero-CBD-F271C conjugate exhibited the largest (ca 3.5-fold) fluorescence change in response to binding of activated cdc42 (see Figure 15).

Fluorescence increase as a measure of CBD-cdc42 interaction - The functionality and specificity of mero-CBD was characterized by measuring fluorescence in the presence of saturating amounts of cdc42-GDP or cdc42-GTPγS. Approximately

- 3.5-fold increase in both excitation and emission maxima was observed, when the probe was bound to cdc42-GTP γ S, but not cdc42-GDP. Negligible increase (<5%) was also observed in the presence of activated Rac1. No effect on the mero-CBD fluorescence was observed when RhoA-GTP γ S was present.
- 5 Furthermore, even in the presence of excess of activated Rac1 and RhoA, GDP- and GTP γ S-bound forms of cdc42 were easily distinguished by CBD-A4C fluorescence (data not shown). This result demonstrated the suitability of the CBD biosensor for use in live cells where different Rho GTPases may be present at comparable concentrations.
- 10 Figure 17 demonstrates use of the biosensor in living cells. To eliminate potential artifacts due to varying cell thickness, uneven illumination, etc., the biosensor was loaded into the cell together with CBD labeled with nonresponsive Alexa546 fluorophore. The ratio of the mero-CBD image to the CBD-Alexa image provided a quantitative measure of the extent of GTPase activation. The lighter, warmer colors show areas of higher cdc42 activation.
- 15

Example 9: Rhodamine-Tagged Peptide Biosensors

Proteins within living cells were fluorescently labeled *in situ*, without isolation or reintroduction of the protein. A short peptide tag derived from the leucine zipper of GCN4 transcription factor was fused to a cellular protein and this fusion protein was expressed in live cells. A second peptide from GCN4, which binds with high affinity to the peptide fused onto the cellular protein, was covalently labeled with rhodamine and also introduced into live cells, where it specifically and selectively labeled the tagged protein. Attachment of rhodamine at the chosen site provided good peptide-peptide binding affinity (5 nM Kd).

20 Two proteins were labeled and visualized in living cells, the cytoplasmic protein α -actinin, and a protein spanning the endoplasmic reticulum (ER) membrane, the α -chain of the high affinity IgE receptor (F ϵ RI). The latter membrane bound

25 protein has previously been inaccessible to synthetic dye attachment through isolation and labeling.

30

Materials and Methods

Peptide synthesis. Peptides were synthesized manually using solid-phase methods with in situ neutralization and HBTU activation procedures for Boc chemistry, on either -OCH₂-Pam or MBHA resins as previously described.

- 5 Standard Boc-protecting groups strategies were employed except for Lys 21 on the labeled peptide, which was protected with an FMOC base-labile protecting group to allow for selective deprotection. Purification was performed using reversed-phase high performance liquid chromatography (RP-HPLC) to obtain peptides greater than 95% pure by analytical HPLC (UV detection at 214 nm and
- 10 linear gradients of solvent B (0.09% TFA in 90% acetonitrile/10% water) in solvent A (0.1% TFA in water)). Labeled peptide expected mass (w/o FMOC) = 3846, observed mass = 3846 +/- 0.7. Protein tag peptide expected mass = 3735, observed mass = 3735 +/- 0.7.
- 15 Derivatization of the labeled peptide at Lys 21 with rhodamine.
The labeled peptide employed had the following sequence.
CEMAQLEKEVQALESEVASLEKEVQALEKEVAQR-NH₂ (SEQ ID NO:15)
This peptide binds with specificity to the following peptide (“tag”) sequence.
KMAQLKKKVQALKSKVASLKKKVQALKKKVAQR-NH₂ (SEQ ID NO:16)
- 20 Lysine 21 of the peptide having SEQ ID NO:15 was selected as the site for tetramethylrhodamine attachment. During Boc-solid phase synthesis, the lysine was protected with a fluorenylmethoxy-carbonyl protecting group, enabling selective deprotection and ready labeling while the peptide was still attached to the synthesis resin. Peptides containing three heptamer repeats were used to
- 25 produce low nanomolar affinities.

To begin the attachment procedure, 20 mg N-hydroxysuccinimide (0.174 mmol) and 30.6 mg of 5-(and 6-) tetramethylrhodamine (0.071 mmol) were dissolved in 300 microliters of DMF. To this mixture, 10.85 microliters of diisopropylcarbodiimide were added. The reaction was then mixed and incubated

- 30 at room temperature for 2 hours. After 2-3 minutes, diisopropylurea precipitation was evident in the activated dye reaction. 200 mg of resin with attached peptide (~0.05 mmol peptide) was treated 3 times with 50% piperidine

in DMF to selectively remove the Fmoc protection of Lys 21 (total piperidine deprotecton time was 15 minutes). Deprotection of Lys 21 was confirmed by quantitative ninhydrin assay. After Fmoc removal, the resin was thoroughly washed with DMF and drained completely. The *in situ* activated

- 5 tetramethylrhodamine N-hydroxysuccinimide ester was diluted to 400 microliters and added to the peptide-bearing resin dropwise until the resin was saturated with activated dye solution. The reaction was incubated at room temperature for 4 hours, washed with DMF and then with DCM/MeOH (1/1) until washes contained minimal color (4-5 washes each). The resin was dried
10 under vacuum overnight. Standard protocols for cleavage/deprotection and purification of tetramethylrhodamine-labeled label peptide were utilized.

Expected mass = 4258, Observed mass = 4258 +/- 0.36.

Peptide *in vitro* interaction measurements. Steady-state fluorescence

- 15 measurements were performed on an SLM 8100 spectro-fluorometer at 25°C with spectral bandwidths set to 8/16/8 nm and 16/16 nm for excitation and emission, respectively. The absorption of the tetramethylrhodamine (TMR) dye linked to the lysine side-chain was used to calculate the solution concentration of the peptide in 10 fold diluted PBS, using the assumption that the extinction
20 coefficient of lysine-linked TMR is the same as that of free dye. The TMR absorption maximum was 553 nm on the peptide vs. 550 for free dye. The extinction coefficient of the free dye was $75916 \text{ M}^{-1}\text{cm}^{-1}$. For these assays only, the sequence YGRKKRRNRRRP (SEQ ID NO:17) was appended to the N-terminus of the rhodamine-containing peptide, as this sequence was being tested
25 in a parallel study of cell import. For the anisotropy titrations, the labeled peptide at a concentration of 12.52 nM was incubated with a stock solution of unlabeled peptide in 0.1x PBS. Titrations were performed at magic angle setting with the excitation wavelength set to 553 nm and the emission wavelength to 580 nm. The spectra were corrected for Raman scattering from the PBS solution, for dilution,
30 and checked for interactions of the label with the ligand by a parallel titration with label peptide using 5-carboxytetramethylrhodamine in PBS in the reference cuvette. For each addition of unlabeled peptide, ten anisotropy measurements

were taken, and the total intensity and anisotropy were calculated according to standard equations with the G-Factor consistency determined at each titration point. The average of the anisotropy was used for data analysis. Standard deviations were <1%. If the quantum yield of the fluorophore changed upon binding, this change had to be taken into account for calculation of the ratio between free and bound biosensor-tag. The anisotropy binding data were therefore fit to the following equation that describes the anisotropy enhancement of tetramethylrhodamine in the labeled peptide (TMRPep) as function of the total concentration of added unlabeled peptide (UPep).

10

$$r = \frac{(r_{\max} * Q - r_{\min}) * [TMRPep / UPep]}{r_{\min} + \frac{[(TMRPep)_0]}{1 - (1 - Q) * [TMRPep / UPep]}}$$

with $Q = q_{\text{bound}}/q_{\text{free}}$, the ratio between quantum yield of TMRPep in bound to free form, and $[TMRPep/UPep] = (TMRPep) + [UPep] + Kd - (([TMRPep] + [UPep] + Kd)^2 - 4 * [TMRPep]_0 * [UPep]_0^{1/2})/2 * [TMRPep]_0$.

Expression Constructs

The various constructs used in this study were similarly prepared using 20 standard cloning (Sambrook et al) of PCR-generated coding sequences into the indicated vectors. The sequence of all constructs was confirmed using an ABI model 377 version 3.0 DNA sequencer.

Cloning pEGFP- α -actinin1-peptide tag. The oligonucleotides,
25 5'GAGCTGTACAAGGGAGG-CATGAAGATGGCCAGCTG AAG 3' (SEQ

ID NO:18) and 5'GTCGCGGCCGCCATTATCGCTGGGCCAC-CTTCTTC3' (SEQ ID NO:19) synthesized by Gibco Life Technologies, were used to amplify the peptide tag insert. The humanized DNA sequence for the peptide tag, derived from the available peptide, was amplified from pCDNA-Nova2-EGFP-5 cdc42 (T17N) to incorporate a BsrGI restriction site on the 5' end and a NotI restriction site on the 3' end. These blunt end PCR products produced by pfu polymerase were ligated into pEGFP-N1- α -actinin. After BsrGI and NotI restriction enzymes were used the construct was gel purified (Qiagen Qiaquick Gel Extraction Kit) and ligated with T4 ligase (Gibco, Life Technologies) into 10 the pEGFP-N1- α -actinin1 vector downstream of EGFP. The pEGFP-N1- α -actinin1-Nova2 sequence was verified through sequencing performed by The Scripps Research Institute core facility.

Cloning of peptide-tagged Fc ϵ RI α -chain. The cDNA encoding the human high 15 affinity IgE receptor (Fc ϵ RI) α -chain containing the NOVA-2 tag positioned at C-terminus was generated by a succession of four PCR reactions to sequentially build the desired full length coding sequence, which was then cloned into the pcDNA3.1(+)zeo vector (Invitrogen, Carlsbad, CA) to sequentially build the 20 desired full length coding sequence. Each amplification used the 5' primer: 5'- GACTGGATCCGAGTCCATGAAGAAGATGGCTCCTGCC (SEQ ID NO:20) together with the 3' primers described below, using human Fc ϵ RI α -chain plasmid as template and Pfu polymerase (Stratagene; San Diego, CA) with 25-30 cycles of 94° C for 1 min; 65° C for 1 min; and 72° C for 2 min, followed by one cycle of 72° C for 10 min. The following downstream PCR primers were used to 25 generate PCR products 1-4, respectively.

- 1: 5'CTGAACCTTCTTCTTGAGCTGGCCATCTTACTTCCGCCACCGTTG
TTTTGGGGTTGGCTTAGG (SEQ ID NO:21);
2: 5'CACCTTTCTTCAAGCTGCCACTTGCTCTGAGCGCCTGAACCT
TTCTTCTTGAGCTGGGC (SEQ ID NO:22);
30 3: 5'GGCCTCGAGTCAACGTTGAGCGACTTCTTTTAAGGCTTGCACC
TTTTCTTCAAGCTGCCAC (SEQ ID NO:23);

4: 5'-CTCGATCGCTCGAGTCAACGTTGAGCGACTTCTT (SEQ ID NO:24).

The full-length PCR product from reaction 4 was isolated and then cloned as a Xho I/Hind 3 fragment into the pcDNA3.1(+) zeo vector (Sambrook, J., E. F.

5 Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.). The nucleotide sequence of the NOVA2-tagged α -chain coding sequence was confirmed on an ABI model 377 version 3.0 DNA sequencer.

10 Cloning of the GFP-MHC HLA B27 Fusion Protein.

The GFP protein was fused to the ER-resident major histocompatibility complex I glycoprotein D_b (MHC). The cDNA encoding the human HLA B27 was generated by PCR amplification using the following oligonucleotide primers: 5'-GGGGATCCTCTCAGACGCCG-3' (SEQ ID NO:25) and 5'-

15 CATGCCATGGCTCCGGATCCACCAGCTGTGAGAGACAC-3' (SEQ ID NO:26) to produce BamHI and NcoI ends. This product was ligated with the GFP coding sequence that had been previously excised from the pSGFP vector as a NcoI-NotI fragment. This ligation product was then cloned into the BamHI-NotI sites of the vector pcDNA3 to produce the desired MHC-GFP
20 fusion protein construct.

Cell culture and imaging. Cos-7 cells were cultured and transfected with FuGENE6 (Roche Molecular Biochemicals) according to the manufacturer's protocol. Cells expressing EGFP- α -actinin were microinjected with 20- μ M
25 Rhodamine-labeled peptide in water. After injection, cells were washed with Dulbecco's phosphate buffered saline (DPBS) (Life Technologies) with 10% FBS, then mounted in a heated chamber on a Zeiss Axiovert 100TV microscope and maintained in DPBS with 10% Fetal Bovine Serum (FBS) to reduce background fluorescence. Images were obtained using a Photometrics PXL
30 cooled CCD camera with 1x1 binning, and a Zeiss Fluar 40 x 1.3 NA oil immersion objective. Fluorescence filters from Chroma were as follows: GFP: HQ480/40, HQ535/50, Q505LP; Rhodamine: HQ545/30, HQ610/75, Q565LP.

Images were background subtracted and contrast stretched using Inovision ISEE software, then formatted for display using Adobe Photoshop software.

Results

- 5 To assess the ability of the peptide to maintain tight binding after labeling with dye, the affinity of the labeled peptide-peptide tag interaction was determined by equilibrium fluorescence titration, monitoring the changes in anisotropy and fluorescence intensity (see Figure 18). High affinity was critical for live cell fluorescence imaging to maximize the proportion of labeled peptide bound to the targeted protein. Unattached labeled peptide could reduce sensitivity by generating uniform fluorescence background, and higher free peptide concentrations could perturb normal cell function or generate appreciable nonspecific binding. Upon binding, the labeled peptide showed a drop in quantum yield and a more than 230% increase in rhodamine anisotropy,
- 10 indicating considerable reduction in rotation of the dye. Fitting the anisotropy values and peptide concentrations to a quadratic equation describing a simple 1:1 binding interaction (see experimental procedures) indicated that the interaction had a K_d value of 5.4 ± 1.1 nM. This high affinity suggested that there was little or no influence of the dye on the peptide-peptide interaction.
- 15 Next, α -actinin in living cells was visualized and the fluorescence of the GFP-fusion protein was compared to that of the peptide labeled with rhodamine. An α -actinin construct with GFP fused to the N-terminus and the tag peptide on the C-terminus was expressed in Cos-7 cells. These cells were loaded with rhodamine peptide through microinjection. The concentration of the rhodamine-labeled peptide was optimized at $20 \mu\text{M}$ to minimize background fluorescence while clearly labeling the tagged peptide. We had determined previously that a uniform fluorescence background is not a serious obstacle using other fluorescently labeled proteins, including rhodamine actin in cytoskeletal fibers, which produces a physiologically normal background of unpolymerized protein.
- 20 Figures 19a and 19c show GFP images and Figures 19b and 19d show rhodamine images taken of the same cells. The fluorescence from a different cell is shown in Figure 20. The peptide tag gave remarkably detailed images of
- 25
- 30

α-actinin distribution, which was difficult to distinguish from those obtained using GFP. Controls in which only the rhodamine-peptide or the α-actinin GFP were present in the cell proved that the colocalization of the rhodamine and GFP signals was not due simply to imperfect selectivity of the fluorescence filters 5 used to separate rhodamine and GFP signals. (See Figure 21.) These controls also demonstrated that binding to other leucine zipper proteins or undesired sites was not a problem, although a slight affinity of rhodamine for the nuclear envelope could be seen.

An important advantage of our labeling methodology is the ability to 10 label proteins that cannot be readily isolated or reintroduced in the cell. This was demonstrated by fusing the tag peptide to the C-terminus of the FcεRI, a membrane-spanning protein that remains in the endoplasmic reticulum unless co-expressed with the FcεRI γ-subunit. The tagged protein was expressed in Cos-7 cells together with a GFP fusion of the ER-resident major 15 histocompatibility complex I glycoprotein D_b (MHC), used here as an ER marker.

Figure 22 shows that the ER-localized FcεRI α-chain was clearly visualized using the rhodamine-tagged peptide. The distribution of the two proteins was not identical, with the Fc receptor showing some concentrations in 20 the perinuclear region. Controls using each fluorescent species alone in cells showed that results were not due to bleedthrough of emission from one fluorophore into the image of the other, and that localizations were not the result of nonspecific binding (data not shown).

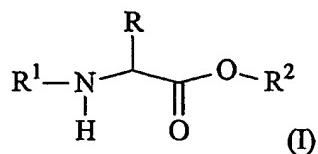
In summary, this approach provides access to many proteins which were 25 previously very difficult to label with synthetic fluorophores for live cell experiments. Unlike other methods requiring relatively difficult synthesis of reagents or bulky antibody tags, this procedure can be accomplished using routine cloning and peptide synthesis procedures.

30 All publications, patents, and patent documents are incorporated by reference herein, as though individually incorporated by reference. The invention has been described with reference to various specific and preferred

embodiments and techniques. However, it should be understood that many variations and modifications can be made while remaining within the spirit and scope of the invention.

WHAT IS CLAIMED IS:

- 1. A compound of formula (I):**

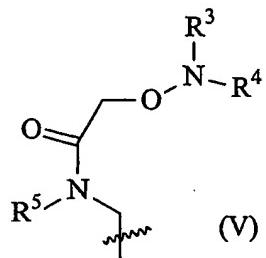


wherein

- 5 R¹ is hydrogen or an amino protecting group;
R² is hydrogen or a carboxy protecting group;
R is an organic radical comprising one or more aminoxy groups.

2. The compound of claim 1 wherein R is a radical of formula (V):

10



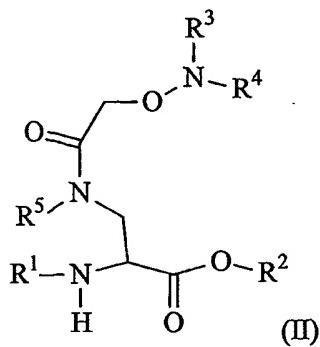
wherein

- R³ is hydrogen, (C₁-C₆) alkyl, an amino protecting group, or a radical comprising one or more aminoxy groups;

15 R⁴ is hydrogen, or an amino protecting group; and

R⁵ is hydrogen, or (C₁-C₆)alkyl.

3. The compound of claim 2 which is a compound of formula (II):



wherein:

- R^1 is hydrogen or an amino protecting group;
- R^2 is hydrogen or a carboxy protecting group;
- R^3 is (C_1-C_6) alkyl;
- R^4 is hydrogen, or an amino protecting group; and
- R^5 is hydrogen.

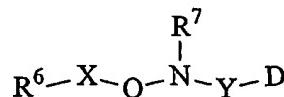
10 4. The compound of claim 3 wherein R^4 is 2-chlorobenzyloxycarbonyl.

5. The compound of claim 1 which is α -benzyloxycarbonyl- β -[N -(2-chlorobenzyloxycarbonyl)- N -methylaminoxyacetyl]- α , β -diaminopropionic acid.

15 6. A peptide comprising a backbone and one or more aminoxy groups; provided the peptide is not glutathione; and provided the peptide has at least one aminoxy group that is not part of a group $H_2N-O-CH_2-C(=O)-$ positioned at the N-terminus of the peptide or that 20 is not part of a group $-C=N-O-CH_2-C(=O)-$ that is in the backbone.

7. A peptide comprising a backbone and one or more secondary aminoxy groups; provided the peptide has at least one aminoxy group that is not part of 25 an oxime ($C=N-O-$) in the backbone.

8. A peptide conjugate of formula (III):



(III)

wherein

R^6 is a peptide or polypeptide;

X is a direct bond or a linking group;

5 R^7 is hydrogen, ($\text{C}_1\text{-C}_6$)alkyl, an amino protecting group, or a radical comprising one or more aminoxy groups;

Y is a direct bond or a linking group; and

D is a functional molecule.

10 9. The peptide conjugate of claim 8 wherein R^6 is an antibody.

10. The peptide conjugate of claim 8 wherein R^6 comprises about 2 to about 1000 amino acids.

15 11. The peptide conjugate of claim 8 wherein R^6 comprises about 5 to about 500 amino acids.

12. The peptide conjugate of claim 8 wherein R^6 comprises about 10 to about 100 amino acids.

20

13. The peptide conjugate of claim 8 wherein X is about 5 angstroms to about 100 angstroms in length.

25

14. The peptide conjugate of claim 8 wherein X is about 5 angstroms to about 25 angstroms in length.

15. The peptide conjugate of claim 8 wherein X is - $\text{R}_a\text{-C}(=\text{O})\text{-NH-R}_b-$ wherein each of R_a and R_b is independently ($\text{C}_1\text{-C}_6$)alkylene.

16. The peptide conjugate of claim 15 wherein each of R_a and R_b is methylene (-CH₂-).
- 5 17. The peptide conjugate of claim 8 wherein R⁷ is hydrogen.
18. The peptide conjugate of claim 8 wherein R⁷ is (C₁-C₆)alkyl.
19. The peptide conjugate of claim 8 wherein R⁷ is methyl.
- 10 20. The peptide conjugate of claim 8 wherein R⁷ is a radical comprising one or more amnooxy groups.
21. The peptide conjugate of claim 8 wherein Y is about 5 angstroms to
15 about 100 angstroms in length.
22. The peptide conjugate of claim 8 wherein Y is about 5 angstroms to about 25 angstroms in length.
- 20 23. The peptide conjugate of claim 8 wherein Y is (C₁-C₆)alkylene.
24. The peptide conjugate of claim 8 wherein Y is methylene (-CH₂-).
25. The peptide conjugate of claim 8 wherein D is a biophysical probe, a peptide, a polynucleotide, or a therapeutic agent.
26. The peptide conjugate of claim 8 wherein D is a cross-linking group or a caged response modifier.
- 30 27. The peptide conjugate of claim 8 wherein D is a biophysical probe.

28. The peptide of claim 27 wherein the biophysical probe is a fluorescent group, a phosphorescent group, a nucleic acid indicator, an ESR probe, a responsive sensor, a caged sensor, or a dye that is sensitive to pH change, ligand binding, or other environmental aspects.

5

29. The peptide of claim 8 wherein D is a peptide.

30. The peptide of claim 8 wherein D is a polynucleotide.

10 31. The peptide of claim 30 wherein the polynucleotide is DNA

32. The peptide of claim 30 wherein the polynucleotide is RNA

33. The peptide of claim 8 wherein D is a therapeutic agent.

15

34. The peptide of claim 8 wherein D is an Alexa dye, a solvatochromic dye, an electrochromatic dye, or a dye that is sensitive to pH change, ligand binding, or other environmental aspects.

20 35. The peptide of claim 8 wherein D is Alexa-532, Hydroxycoumarin, Aminocoumarin, Methoxycoumarin, Amino methylcoumarin, Cascade Blue, Lucifer Yellow, NBD, P-Phycoerythrin, R-Phycoerythrin, (PE), PE-Cy5 conjugates, PE-Cy7 conjugates, Red 613, Fluorescein, BODIPY-FL, BODIPY TR, BODIPY TMR, Cy3, TRITC, X-Rhodamine, Lissamine Rhodamine B,

25 PerCP, Texas Red, Cy5, Cy7, Allophycocyanin (APC), TruRed, APC-Cy7 conjugates, Oregon Green, Tetramethylrhodamine, Dansyl, Indo-1, Fura-2, FM 1-43, DilC18(3), Carboxy-SNARF-1, NBD, Indo-1, Fluo-3, DCFH, DHR, SNARF, or Monochlorobimane, Calcein.

30 36. The peptide of claim 8 wherein D is YOYO-1, Propidium Iodide, Hoechst 33342, DAPI, Hoerchst 33258, SYTOX Blue, Chromomycin A3, Mithramycin, SYTOX Green, SYTX Orange, Ethidium Bromide, 7-AAD,

Acridine Orange, TOTO-1, TO-PRO-1, Thiazole Orange, Propidium Iodide, TOTO-3, TO-PRO-3, or LDS 751.

37. The peptide of claim 8 wherein:

5 R⁶ is (SEQ ID NO. 1);

X is R_a-C(=O)CH(NH₂)CH₂N(H)C(=O)CH₂-R_b; wherein R_a is a direct bond to the amino terminus of R⁶ and wherein R_b is a direct bond to the oxygen of the aminoxy group of formula (III);

R⁷ is methyl;

10 Y is a direct bond; and

D is Alexa-532.

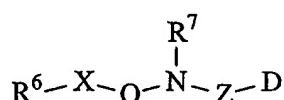
38. A method for preparing a peptide conjugate comprising a peptide linked to a functional molecule, comprising reacting a peptide having one or 15 more secondary aminoxy groups with a corresponding functional molecule having an electrophilic moiety that is reactive with the aminoxy group(s), to provide the peptide conjugate.

39. A method for preparing a peptide conjugate comprising a peptide 20 linked to a functional molecule, comprising reacting a peptide having one or more aminoxy groups with a corresponding functional molecule having an electrophilic moiety that is reactive with the aminoxy group(s), to provide the peptide conjugate; provided that the functional molecule is not a peptide.

25 40. A method for preparing a peptide conjugate comprising a peptide linked to a functional molecule, comprising reacting a peptide having one or more aminoxy groups with a corresponding functional molecule having an electrophilic moiety that is reactive with the aminoxy group(s), to provide the peptide conjugate; provided that the peptide and the functional molecule are not 30 attached through an oxime (C=N-O-) linkage.

41. A method for preparing a peptide conjugate comprising a functional molecule linked to a peptide having a backbone, comprising reacting a peptide having one or more aminoxy groups with a corresponding functional molecule having an electrophilic moiety that is reactive with the aminoxy group(s), to
5 provide the peptide conjugate; provided that when the functional molecule is a second peptide, the functional molecule and the first peptide are not linked through a -C=N-O-CH₂-C(=O)- linkage in the backbone of the first peptide.

42. A peptide conjugate of formula (IV):



(IV)

10 wherein

R⁶ is a peptide, polypeptide or antibody;

X is a direct bond or a linking group;

R⁷ is hydrogen, (C₁-C₆)alkyl, an amino protecting group, or a radical comprising one or more aminoxy groups;

15 Z is a linking group, or Z is a direct single bond or a double bond between N and D; and

D is a functional molecule;

provided that when D is a peptide, N and D are not linked through a -C=N-O-CH₂-C(=O)- linkage in the backbone of the peptide conjugate.

20

43. The peptide conjugate of claim 42 wherein R⁶ is an antibody.

44. A method of identifying an optimal position for placement of a functional molecule on a peptide having a peptide backbone and a known activity, which comprises making a series of peptide conjugates, each peptide conjugate having the same amino acid sequence and the same functional molecule, wherein the functional molecule is linked at a different location along the backbone of every peptide conjugate in the series, and observing which

functional molecule location does not substantially interfere with the known activity of the peptide.

45. The method of claim 44 wherein each peptide of said series of peptide conjugates comprises the peptide conjugate of claim 8 or 42.

46. A method of identifying an optimal position for placement of a functional molecule in a polypeptide having a known activity and an identified peptide segment for attachment of the functional molecule, which comprises:

10 a) making a series of peptide conjugates, each peptide conjugate having the amino acid sequence of the identified peptide segment and the same functional molecule, wherein the functional molecule is linked at a different location along the backbone of every peptide conjugate in the series;

15 b) placing a peptide conjugate within, or at the end of, each polypeptide of a series of polypeptides to create a series of polypeptide conjugates each having the functional molecule at a different location; and

20 c) observing which functional molecule location does not substantially interfere with the known activity of the polypeptide.

47. The method of claim 46 wherein each peptide of said series of peptides comprises the peptide conjugate of claim 8 or 42.

25 48. The method of claim 46 wherein each peptide conjugate is placed within, or at the end of, each polypeptide by natural chemical ligation, intein-mediated protein ligation or chemical ligation.

30 49. A method of identifying an optimal position for placement of an environmentally-sensitive functional molecule on a peptide biosensor having a backbone, which comprises making a series of peptide conjugates, each peptide conjugate having the same amino acid sequence and the same functional

molecule, wherein the functional molecule is at a different location along the backbone of every peptide conjugate in the series, and observing which functional molecule location provides the strongest signal change in response to an environmental change in the peptide conjugate.

5

50. The method of claim 49 wherein each peptide conjugate of said series of peptide conjugates comprises the peptide conjugate of claim 8 or 42.

51. The method of claim 49 wherein each peptide conjugate is part of a
10 polypeptide, antibody, antibody fragment, antibody fragment with linked chains or antibody with linked chains.

52. The method of claim 49 wherein said signal change is a change in phosphorescence, fluorescence emission intensity, fluorescence lifetime,
15 fluorescence excitation wavelength or fluorescence emission wavelength.

53. The method of claim 49 wherein said environmental change in said peptide conjugate is interaction with a target.

20 54. The method of claim 53 which further comprises observing the binding affinity of each peptide conjugate for target or the binding selectivity of each peptide conjugate for target.

25 55. A method of identifying an optimal position for placement of an environmentally-sensitive functional molecule in a polypeptide having a known activity and an identified peptide segment for attachment of the functional molecule, which comprises:

30 a) making a series of peptide conjugates, each peptide conjugate having the amino acid sequence of the identified peptide segment and the same environmentally-sensitive functional molecule, wherein the environmentally-sensitive functional molecule is linked

at a different location along the backbone of every peptide conjugate in the series;

- 5 b) placing a peptide conjugate within, or at the end of, each polypeptide of a series of polypeptides to create a series of polypeptide conjugates each having the environmentally-sensitive functional molecule at a different location; and
- c) observing which functional molecule location provides the strongest signal change in response to an environmental change in the polypeptide conjugate.

10

56. The method of claim 55 wherein each peptide conjugate of said series of peptide conjugates comprises the peptide conjugate of claim 8 or 42.

57. The method of claim 55 wherein each peptide conjugate is placed
15 within, or at the end of, each polypeptide by natural chemical ligation, intein-mediated protein ligation or chemical ligation.

58. The method of claim 55 wherein said signal change is a change in phosphorescence, fluorescence emission intensity, fluorescence lifetime,
20 fluorescence excitation wavelength or fluorescence emission wavelength.

59. The method of claim 55 wherein said environmental change in said protein conjugate is interaction with a target or a change in activation state of a target.

25

60. The method of claim 59 which further comprises observing the binding affinity of each peptide conjugate for target or the binding selectivity of each peptide conjugate for target.

30 61. A polypeptide biosensor which comprises the peptide conjugate of claim 8 or 42.

62. The polypeptide biosensor of claim 61 which comprises a MAP a p21 activated kinase peptide capable of binding Rac, a Wiscott-Aldrich Syndrome Protein peptide capable of binding cdc42, or a portion of a leucine zipper capable of binding another portion of a leucine zipper.

5

63. A polypeptide biosensor comprising a polypeptide capable of binding a GTP-activated Rho GTPase protein, wherein said polypeptide is operatively linked to a functional molecule

10 64. The polypeptide biosensor of claim 63 wherein said Rho GTPase protein is Rac or cdc42.

65. The polypeptide biosensor of claim 63 wherein said polypeptide comprises the protein binding domain of p21 activated kinase 1.

15

66. The protein biosensor of claim 63 wherein said GTP-activated Rho GTPase protein is attached to Green Fluorescence Protein (GFP), Cyan Fluorescence Protein(CFP), Red Fluorescence Protein (RFP) or enhanced GFP (EGFP).

20

67. A fusion protein comprising a biologically active Rho GTPase protein domain operatively linked to a fluorescent protein via the peptide conjugate of claim 8 or 42, wherein said Rho GTPase protein domain is capable of binding GTP and forming an activated GTPase:GTP complex.

25

68. The fusion protein of claim 67 wherein said Rho GTPase protein domain is derived from a Rac, Rho or cdc42 protein.

30 69. The fusion protein of claim 67 wherein said fluorescence protein is green fluorescence protein, cyan fluorescence protein, red fluorescence protein, yellow fluorescence protein, enhanced green fluorescence protein or enhanced yellow fluorescence protein.

70. A nucleic acid encoding the fusion protein of claim 67.
71. An expression vector capable of expressing the fusion protein encoded by the nucleic acid of claim 67.
 - 5
72. A cell comprising the expression vector of claim 71.
73. A method for detecting GTP activation of a Rho GTPase protein in a live cell comprising:
 - 10
 - a) introducing a Rho GTPase protein to said cell, wherein said Rho GTPase protein is operatively linked to a fluorescence dye capable of undergoing fluorescence resonance energy transfer;
 - b) introducing a polypeptide biosensor into said cell, wherein said 15 polypeptide biosensor comprises a polypeptide capable of binding a GTP-activated Rho GTPase protein, and wherein said polypeptide is operatively linked to a fluorescent dye which can undergo fluorescence resonance energy transfer with said fluorescence dye on said GTP-activated Rho GTPase protein when said polypeptide biosensor interacts with said GTP-activated Rho GTPase; and
 - c) observing fluorescence emissions from said polypeptide biosensor within said live cell.
- 25 74. The method of claim 73 which further comprises observing fluorescence emissions from said fluorescence dye on said GTP-activated Rho GTPase.
- 30 75. The method of claim 73 wherein said polypeptide biosensor comprises the peptide conjugate of claim 8 or 42.

76. A method for detecting GTP activation of a Rho GTPase protein, which comprises:

- a) contacting a polypeptide biosensor with a test substance, wherein said polypeptide biosensor comprises a polypeptide capable of binding a GTP-activated Rho GTPase protein, and wherein said polypeptide is operatively linked to an environmentally sensitive dye; and

- b) observing a signal from said polypeptide;

wherein said environmentally sensitive dye will emit a signal of a different lifetime, intensity or wavelength when said polypeptide biosensor is bound to

- 10 said GTP-activated Rho GTPase protein than when said polypeptide biosensor is not bound.

77. A method for detecting GTP activation of a Rho GTPase protein in a live

- 15 cell comprising:

- a) introducing a polypeptide biosensor into said cell, wherein said polypeptide biosensor comprises a polypeptide capable of binding a GTP-activated Rho GTPase protein, and wherein said polypeptide is operatively linked to an environmentally sensitive dye; and

- 20 b) observing a signal from said environmentally sensitive dye within said live cell;

wherein said environmentally sensitive dye will emit a signal of a different lifetime, intensity or wavelength when said polypeptide biosensor is bound to

said GTP-activated Rho GTPase protein than when said polypeptide biosensor is

- 25 not bound.

78. The method of claim 76 or 77 wherein said polypeptide biosensor

comprises the peptide conjugate of claim 8 or 42.

- 30 79. The method of claim 76 or 77 wherein said polypeptide biosensor comprises the protein binding domain of p21 protein kinase or the protein binding domain of Wiscott Aldrich syndrome protein.

80. The method of claim 76 or 77 wherein said Rho GTPase protein domain is derived from a Rac, Rho or cdc42 protein.

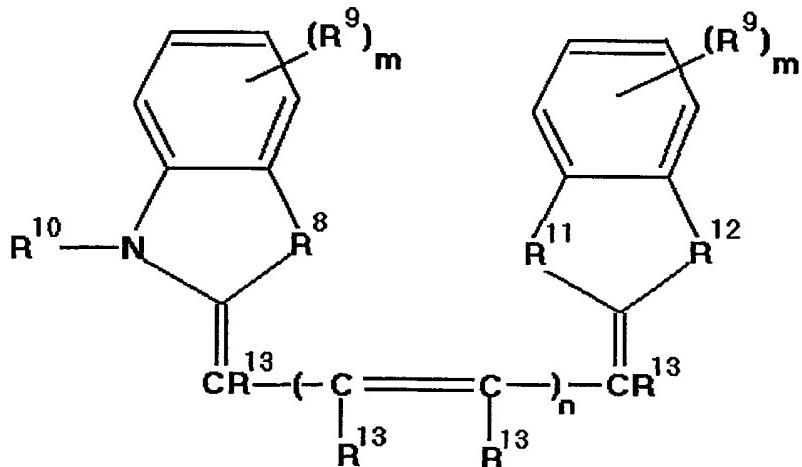
5 81. The method of claim 76 or 77 which further comprises quantifying the amount of GTP-activated Rho GTPase protein.

82. A method for detecting binding of an antibody to an antigen which comprises reacting an antibody comprising the peptide conjugate of claim 8 or
10 42 with an antigen and detecting an antibody-antigen complex.

83. A method for detecting binding of an antigen to an antibody which comprises reacting an antigen comprising the peptide conjugate of claim 8 or 42 with an antibody and detecting an antibody-antigen complex.

15

84. A fluorescent compound of the formula:



wherein:

20 each m is separately an integer ranging from 1-3;
n is an integer ranging from 0 to 5;

R^8 , R^{11} and R^{12} are separately CO, SO₂, C=C(CN)₂, S, O or C(CH₃)₂;

each R^{13} is alkyl, branched alkyl or heterocyclic ring derivatized with charged groups to enhance water solubility and enhance photostability;

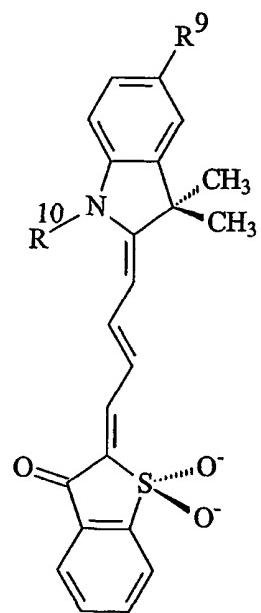
5 each R^9 and R^{10} is separately an alkyl chain derivatized with charged groups to enhance water solubility or with reactive groups for conjugation to other molecules.

85. The compound of claim 84 wherein each R^9 and R^{10} is separately
10 sulfonate, amide or ether.

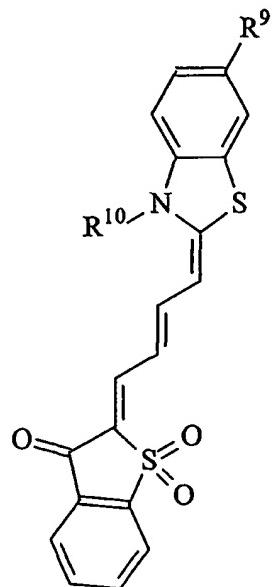
86. The compound of claim 84 wherein each R^9 and R^{10} is separately -NH-(C=O)-CH₂-halide, amine, maleimide, -N=C=O, -N=C=S, acyl halide, succinimidyl ester, sulfosuccinimidyl ester, sulfonyl halide, sulfonyl azide,
15 alcohol, thiol, semicarbazide, hydrazine or hydroxylamine.

87. The compound of claim 84 wherein each R^9 and R^{10} is separately carboxylic acid, alkali or alkaline earth metal salt of carboxylic acid, carboxylic acid activated by carbodiimide, acyl chloride, succinimidyl, sulfosuccinimidyl
20 ester or COORx, wherein Rx is phenol or naphtol further substituted by at least one strong electron withdrawing group.

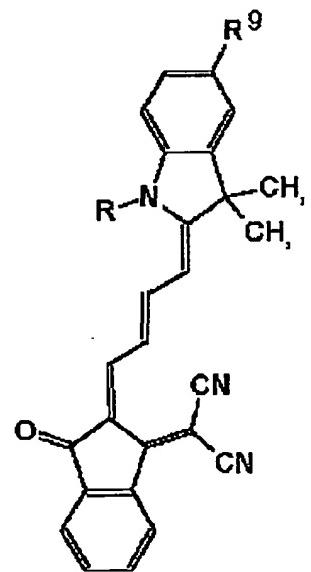
88. The compound of claim 84 having the formula:



89. The compound of claim 84 having the formula:



90. The compound of claim 84 having the formula:



5

91. A peptide biosensor comprising the compound of claim 84.

92. A protein, polypeptide, peptide, antibody, antibody fragment or nucleic acid linked to the compound of claim 84.

93. A method of detecting the location of a cellular protein within a living cell which comprises:

- a. providing the living cell with a biosensor capable of binding to a tag on the cellular protein; and
- b. detecting the location of a functional molecule on the biosensor within the living cell;

wherein said tag is a peptide segment that has been fused to the cellular protein.

94. A method of detecting the location of a cellular protein within a living cell which comprises:

- a. providing the living cell with a biosensor capable of binding to a tag on the cellular protein; and
- b. detecting the location of a functional molecule on the biosensor within the living cell;

wherein said tag is a peptide segment that has been fused to the cellular protein and said biosensor comprises the peptide-conjugate of claim 8 or 42.

95. A method of detecting the location of a cellular protein within a living cell which comprises:

- a. providing the living cell with a biosensor capable of binding to a tag on the cellular protein; and
- b. detecting the location of a functional molecule on the biosensor within the living cell;

wherein said tag is a peptide segment that has been fused to the cellular protein, and said functional molecule is a fluorescent compound of claim 84.

96. The method of claim 93, 94 or 95 wherein said tag has SEQ ID NO:16.
97. The method of claim 93, 94 or 95 wherein said functional molecule is
5 the compound of claim 85, 86, 87, 88, 89 or 90.
98. The method of claim 93, 94 or 95 wherein said cellular protein is
calmodulin, Rho GTPase, rac, cdc42, mitogen-activated protein
kinase, Erk1, Erk2, Erk3, Erk4, IgE receptor (F_cεRI) actin, α-actinin,
10 myosin, or a major histocompatibility protein.
99. A method of attaching a biosensor to a cellular protein within a living
cell which comprises, providing the living cell with a biosensor
capable of binding to a tag on the cellular protein, wherein said tag is
15 a peptide segment that has been fused to the cellular protein
expressed by the living cell.
100. The method of claim 99 wherein said biosensor comprises the
peptide-conjugate of claim 8 or 42.
20
101. The method of claim 100 wherein said peptide-conjugate comprises a
fluorescent compound of claim 84.
102. A nucleic acid encoding the tag fused to the cellular protein of claim
25 99.
103. An isolated vector comprising the nucleic acid of claim 102.
104. An expression vector capable of expressing the tag fused to the
30 cellular protein of claim 99.
105. A cell comprising the expression vector of claim 104.

THIS PAGE BLANK (USPTO)

1/26

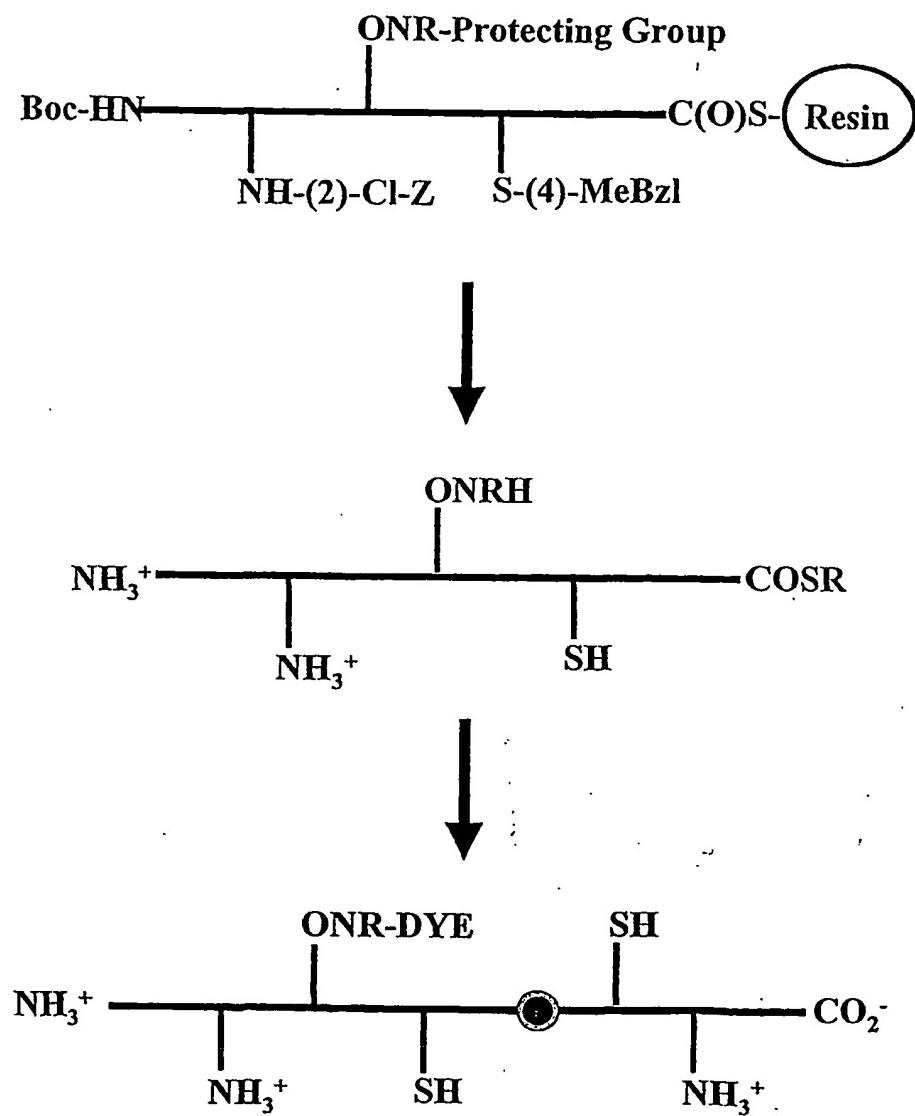


FIG. 1

2/26

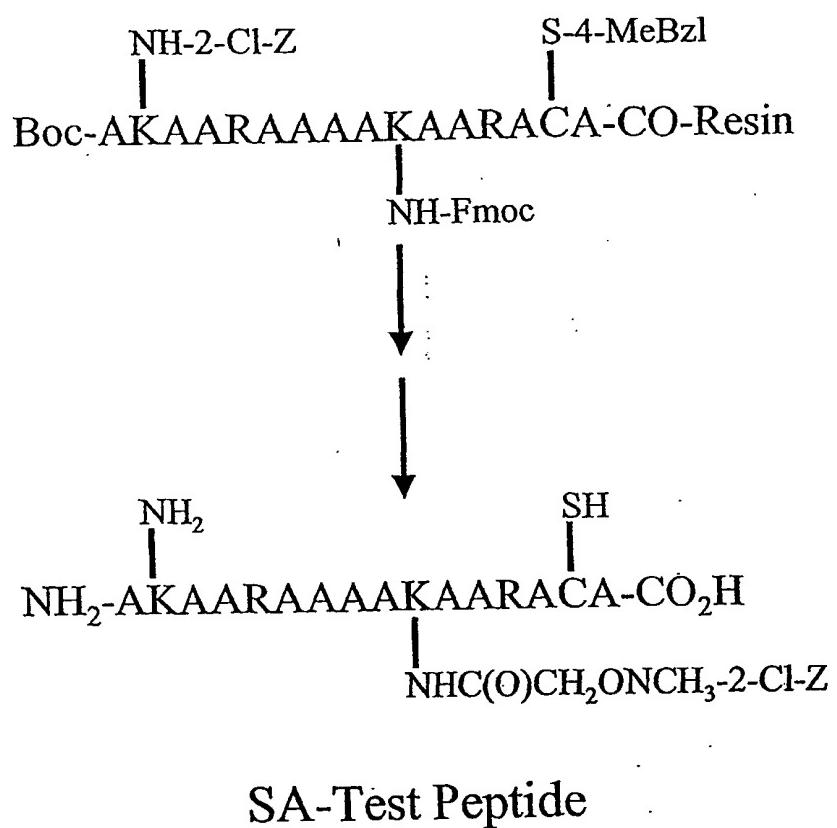


FIG. 2

3/26

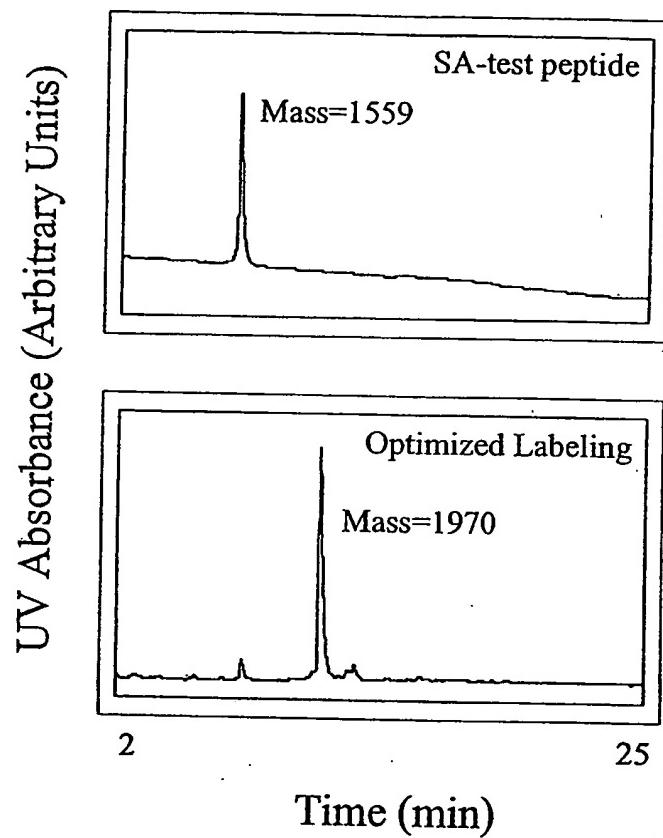


FIG. 3

4/26

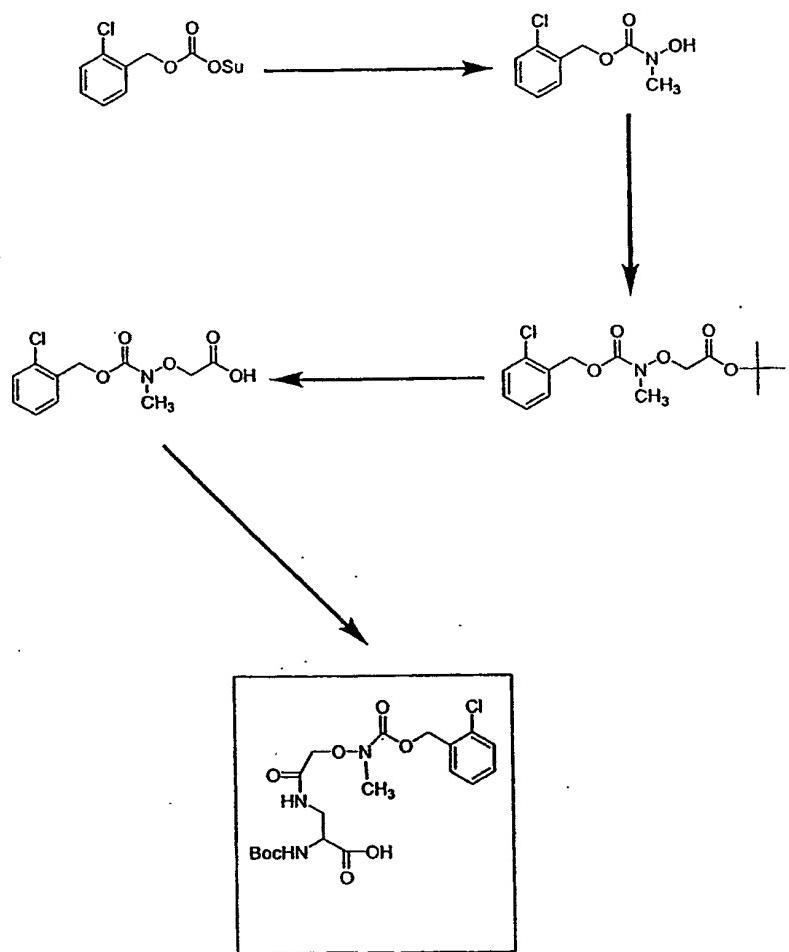


FIG. 4

5/26

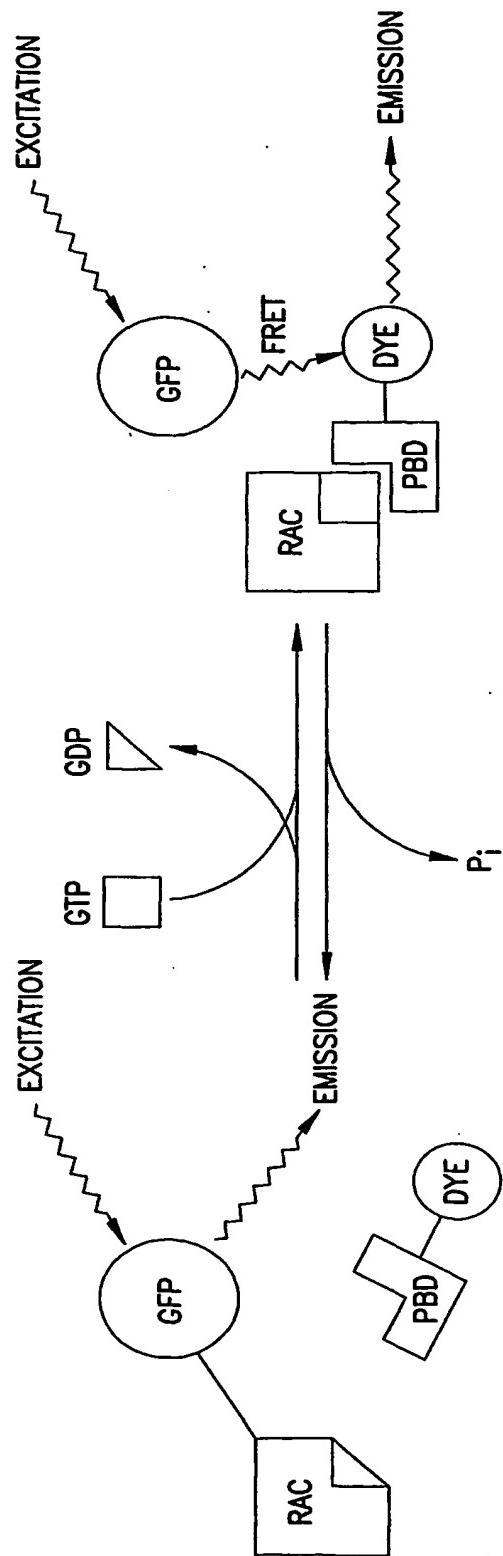


FIG. 5

6/26

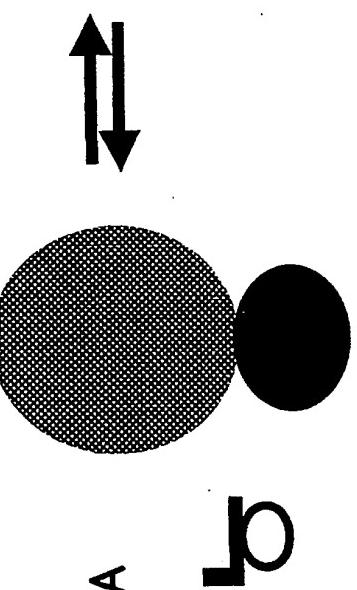
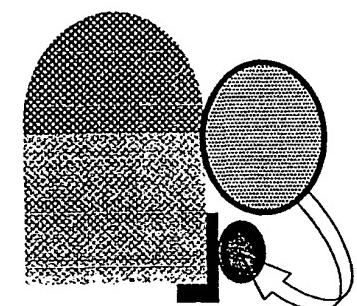
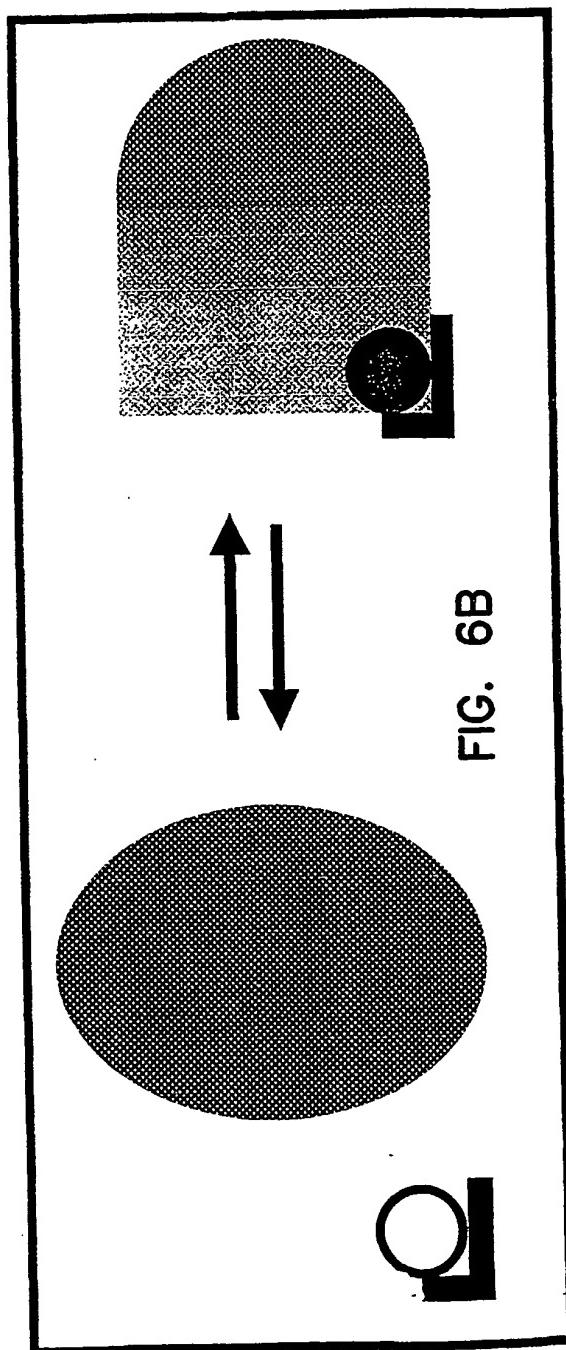


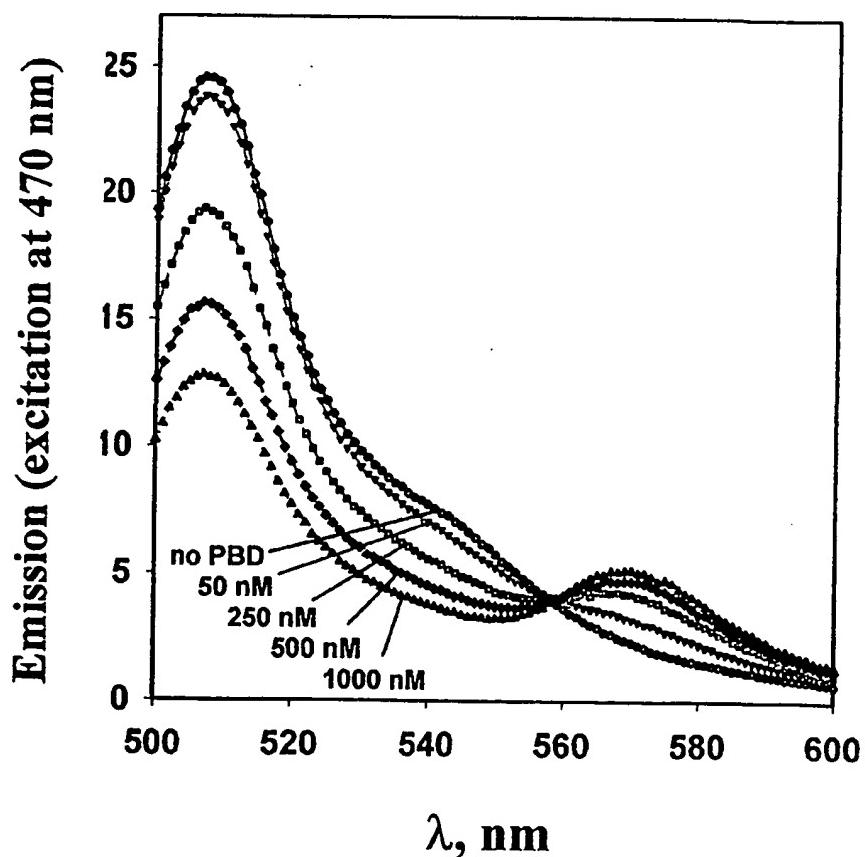
FIG. 6A



BEST AVAILABLE COPY

7/26

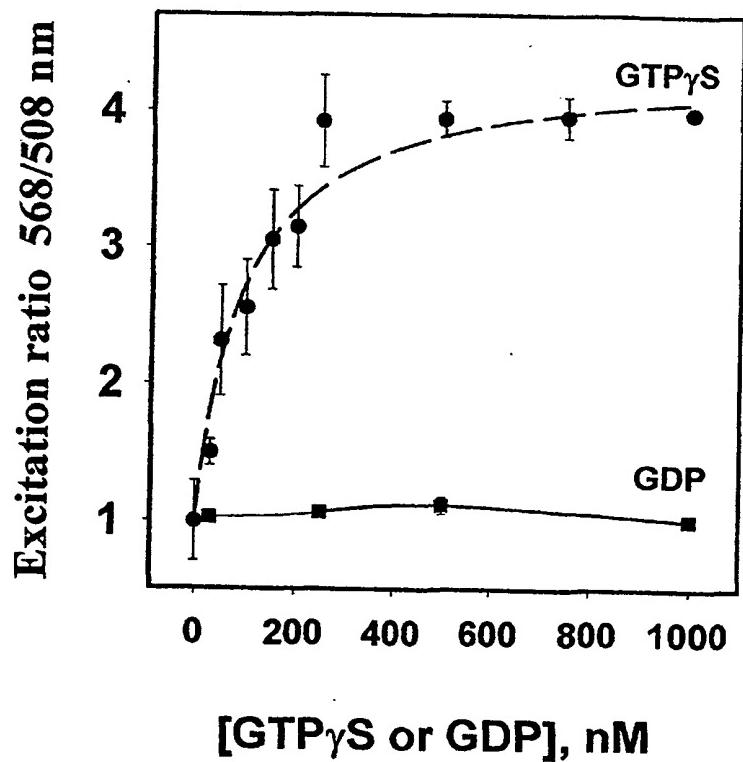
Figure 7A: GFP-Rac to Alexa-PBD FRET



DIS AVAILABLE COPY

8/26

Figure 7B: FRET response to nucleotide state of Rac-GFP



9/26

Individual cells scored for Rac-induced ruffling

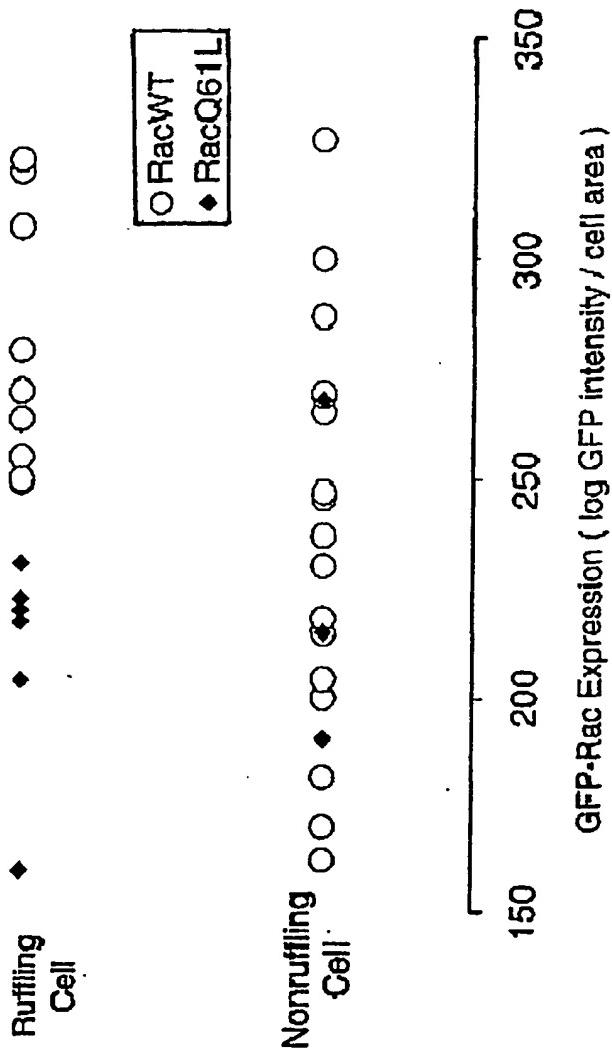


FIG. 8A

10/26

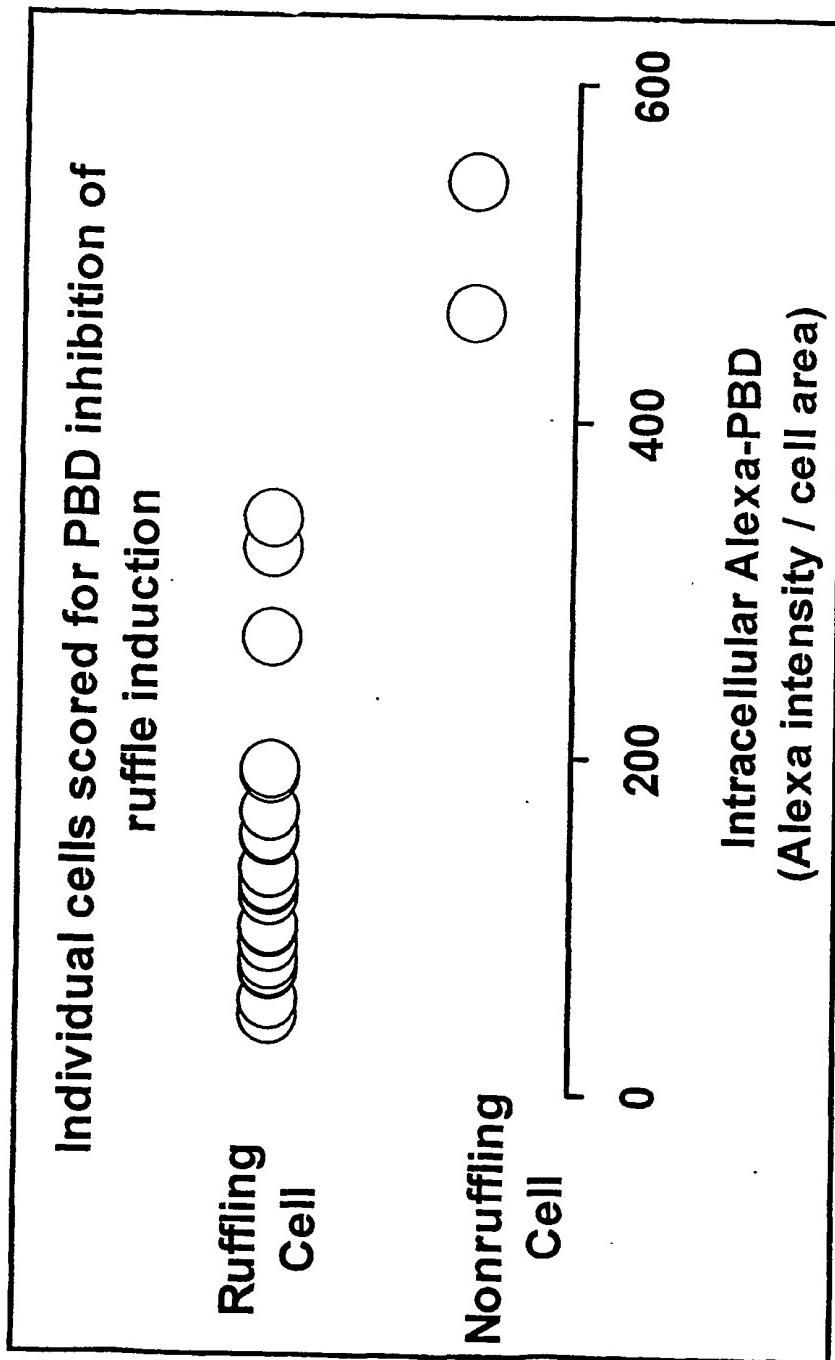
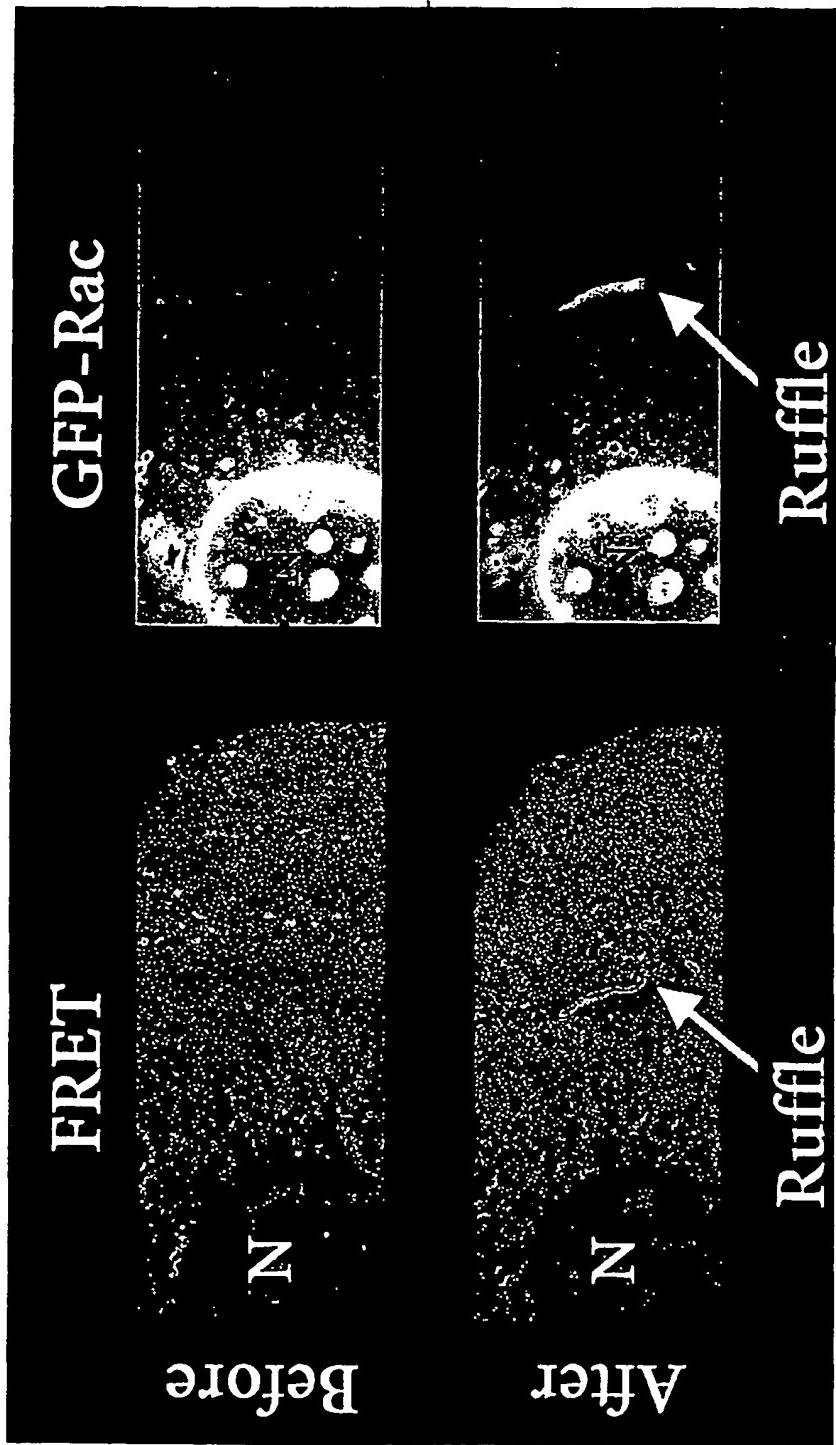


FIG. 8B

11/26

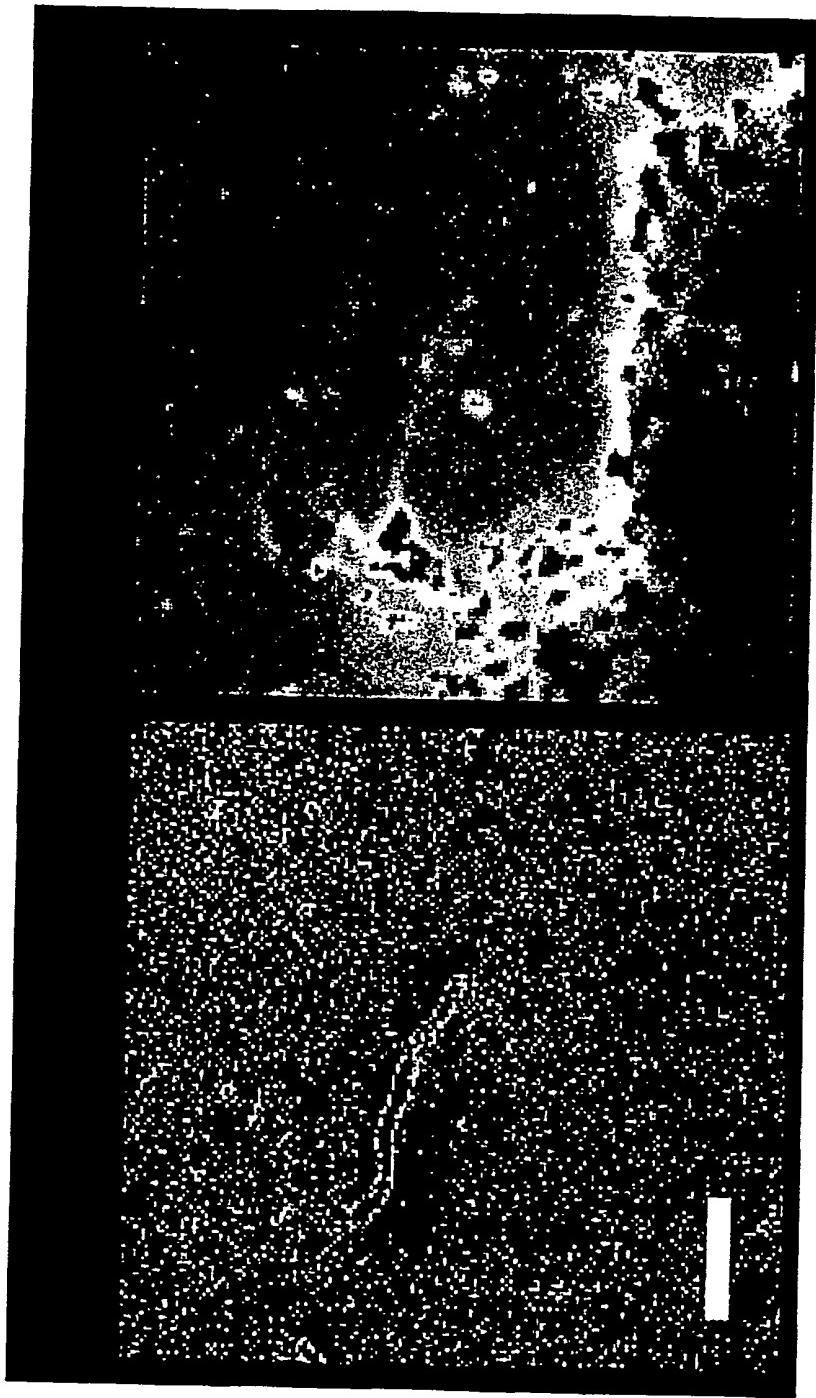
Fig. 9A and 9B: Serum stimulation of a Swiss 3T3 fibroblast



BEST AVAILABLE COPY

12/26

Fig. 9C and 9D: The same ruffle visualized using either FRET or Alexa-PBD localization:



BEST AVAILABLE COPY

Fig. 10A: Rac-GFP

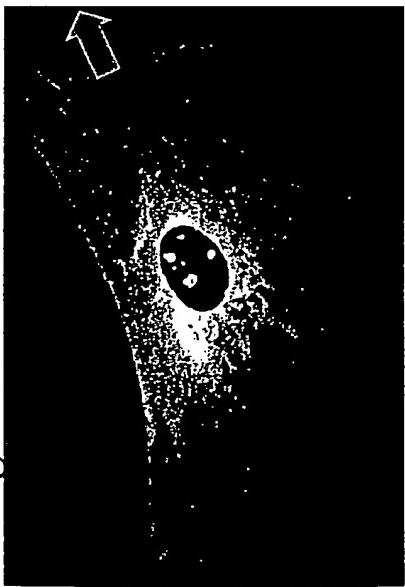
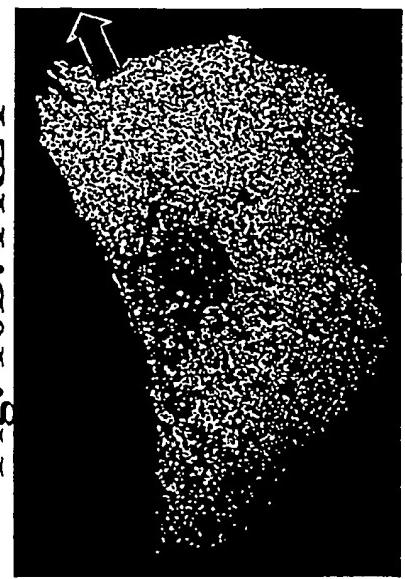


Fig. 10B. FRET



Wound healing

monolayer

Confluent

Magnitude of gradient when highest at front

128 +/- 51 % n=12

Magnitude of gradient highest at rear

9 +/- 4 % n=4

14/26

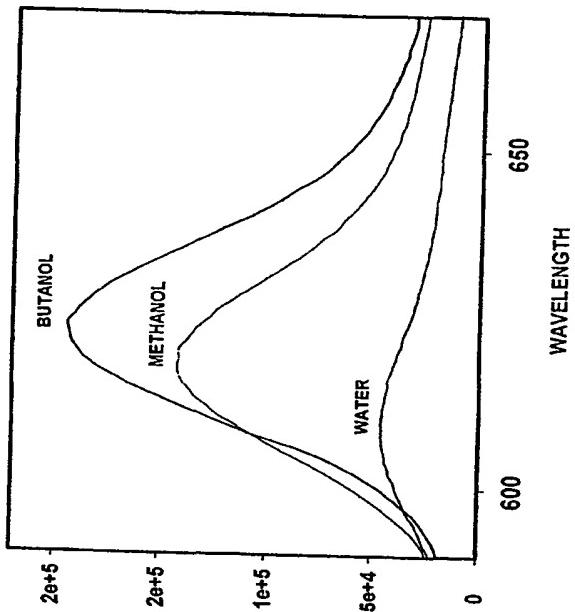
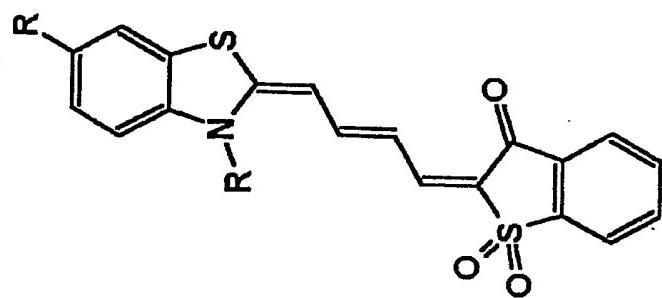


FIG. 11



15/26

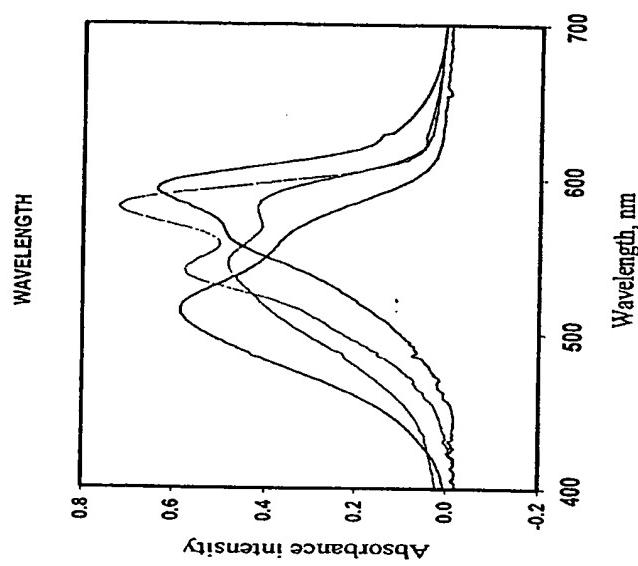
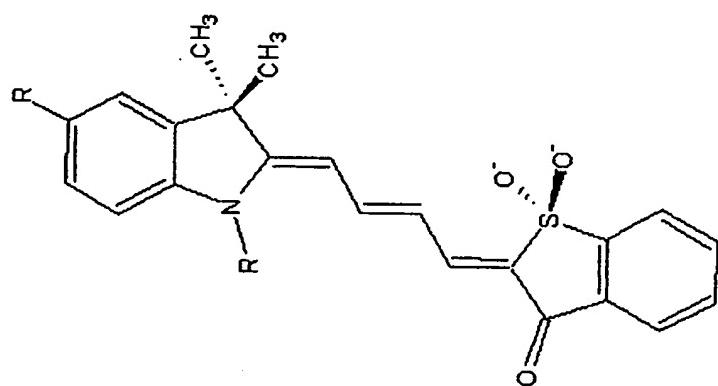


FIG. 12



16/26

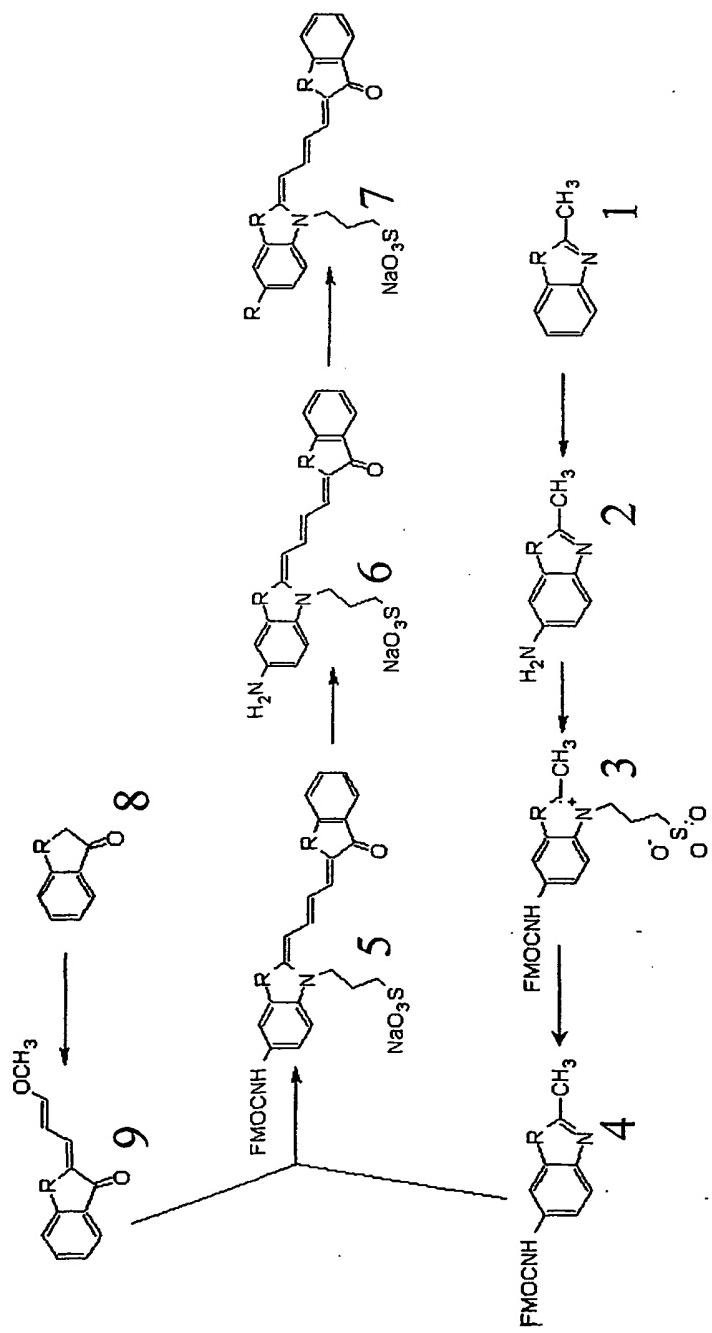


FIG. 13

17/26

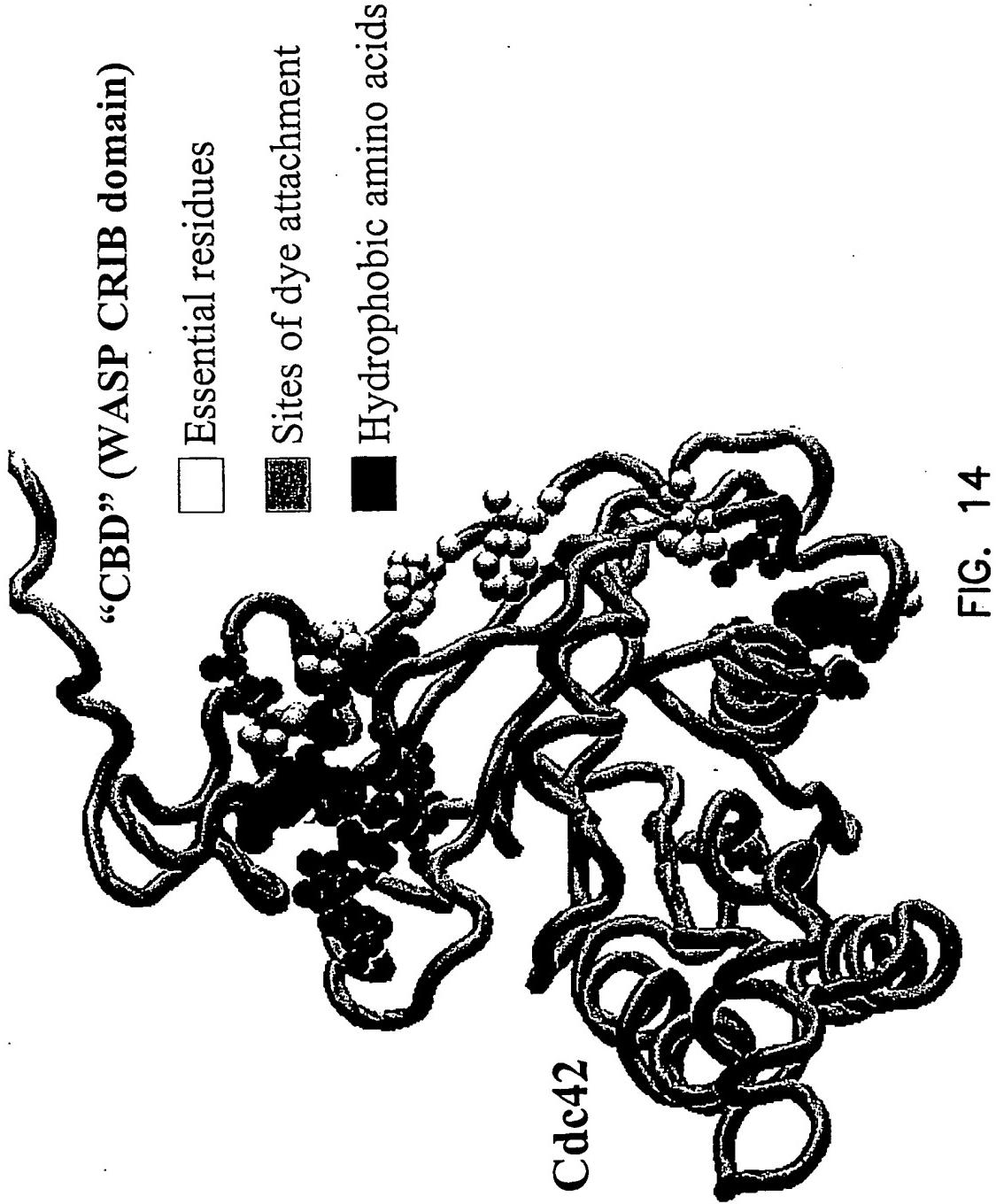
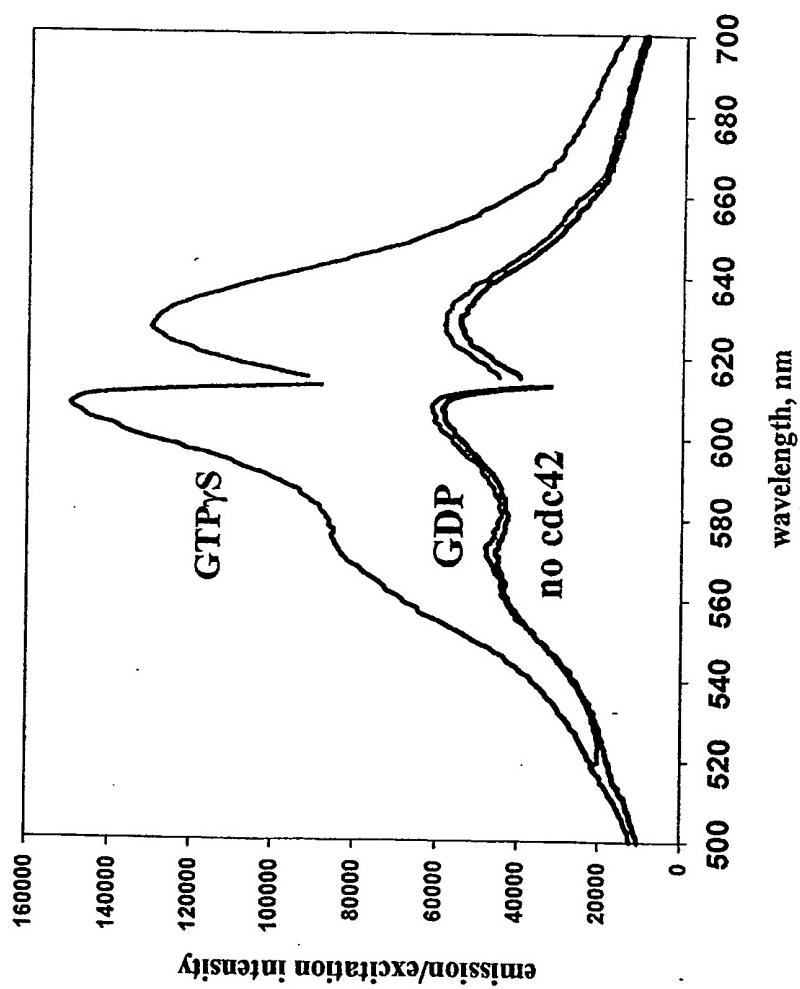


FIG. 14

BEST AVAILABLE COPY

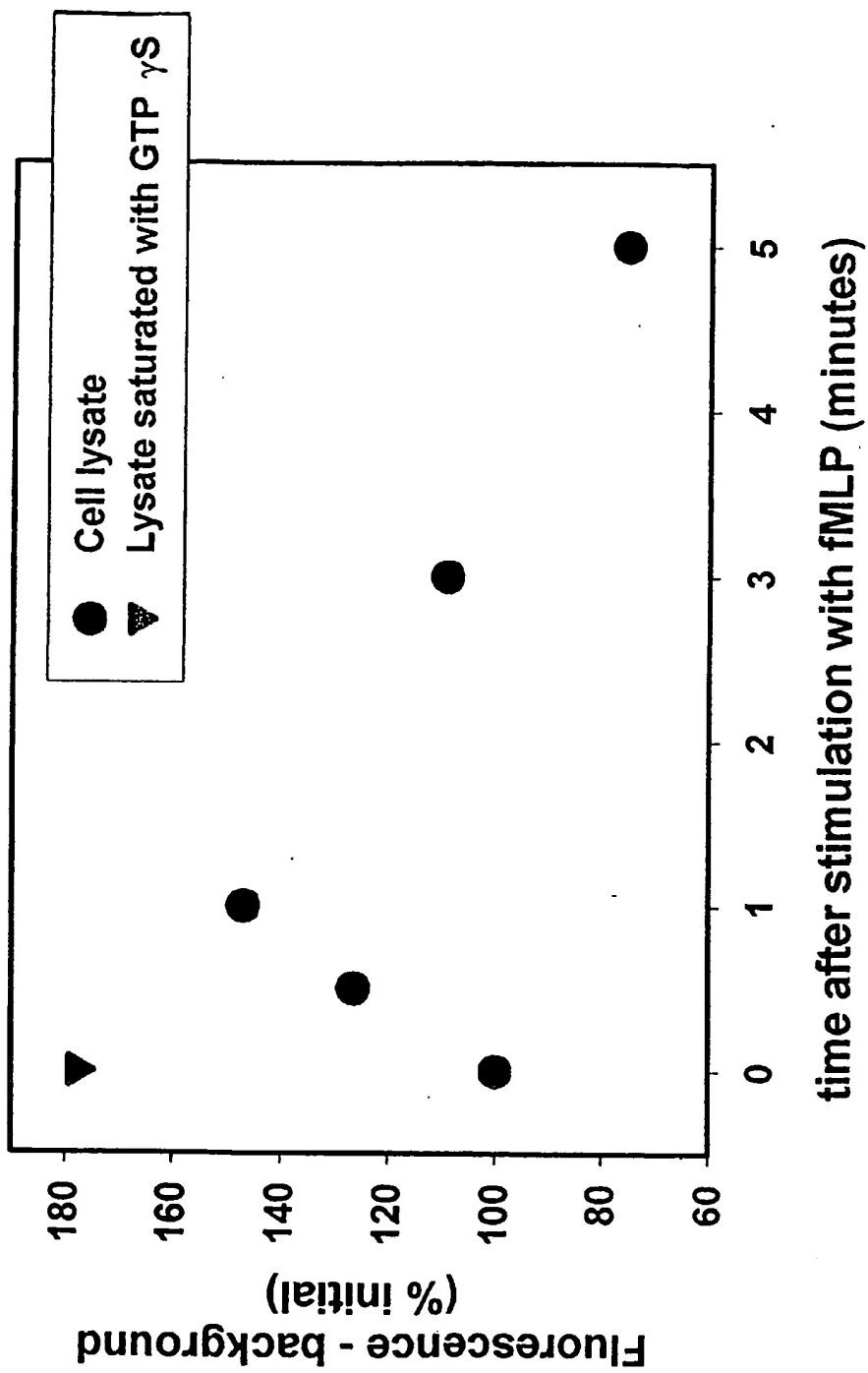
18/26

Fig. 15: Fluorescence of Mero-CBD responds to Cdc42 binding

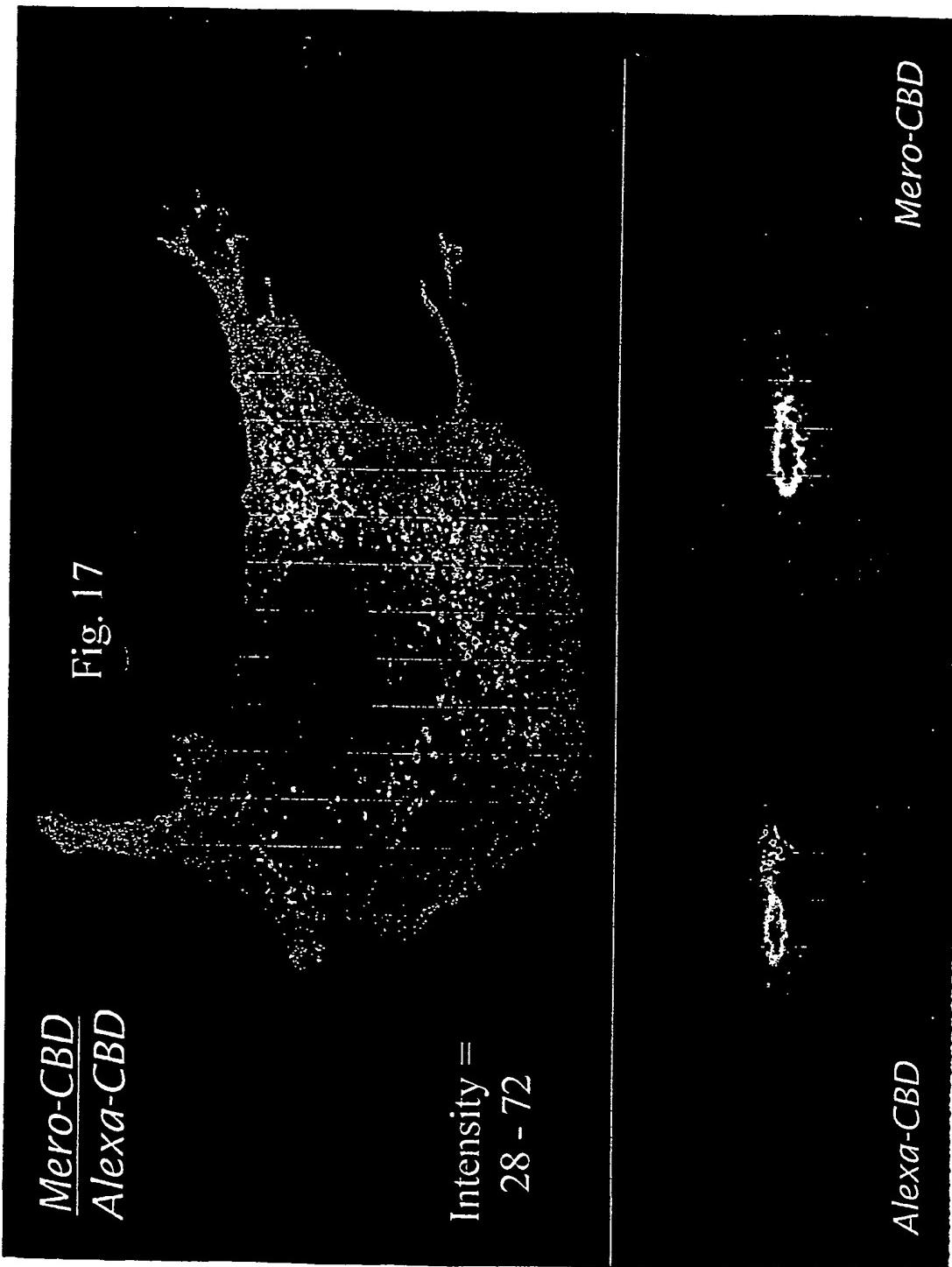


19/26

Fig. 16: Mero-CBD in neutrophil lysates



20/26



BEST AVAILABLE COPY

21/26

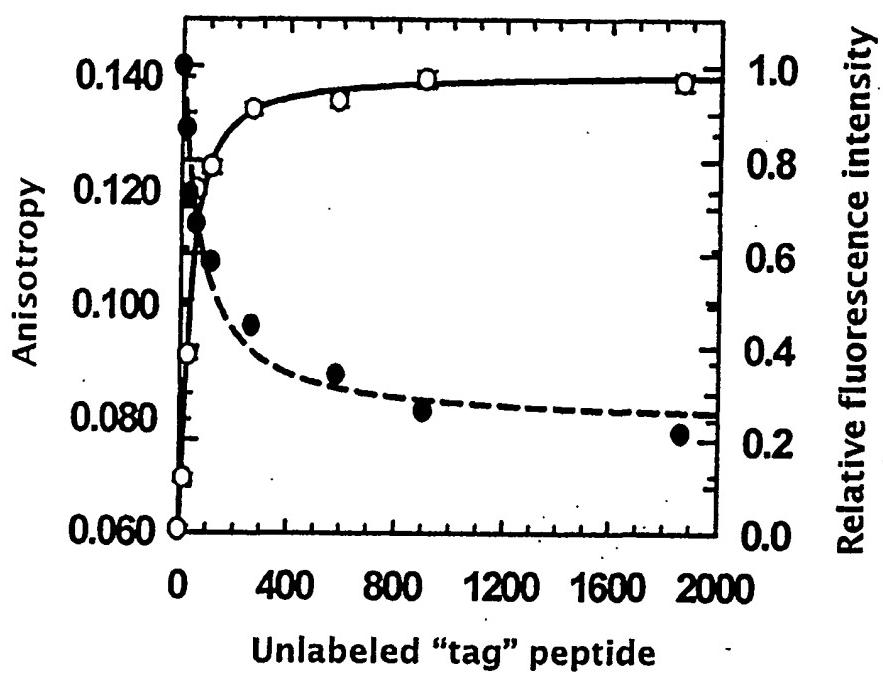


FIG. 18

22/26

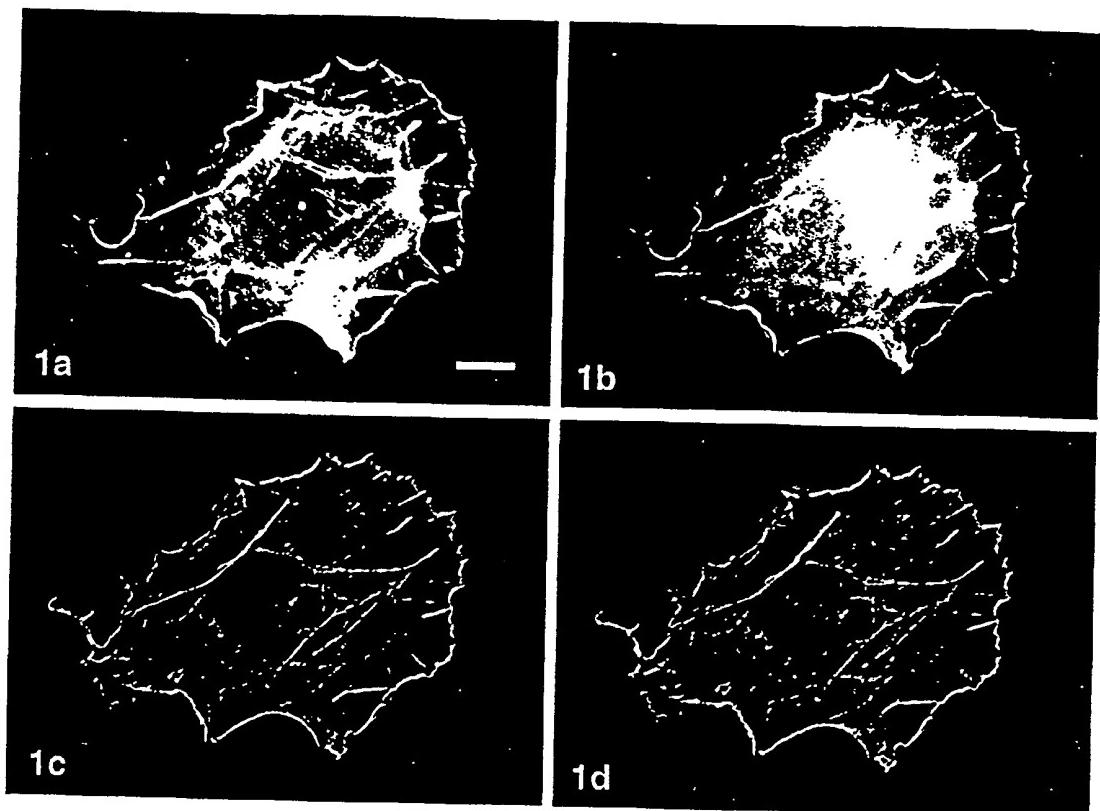


FIG. 19

BEST AVAILABLE COPY

23/26

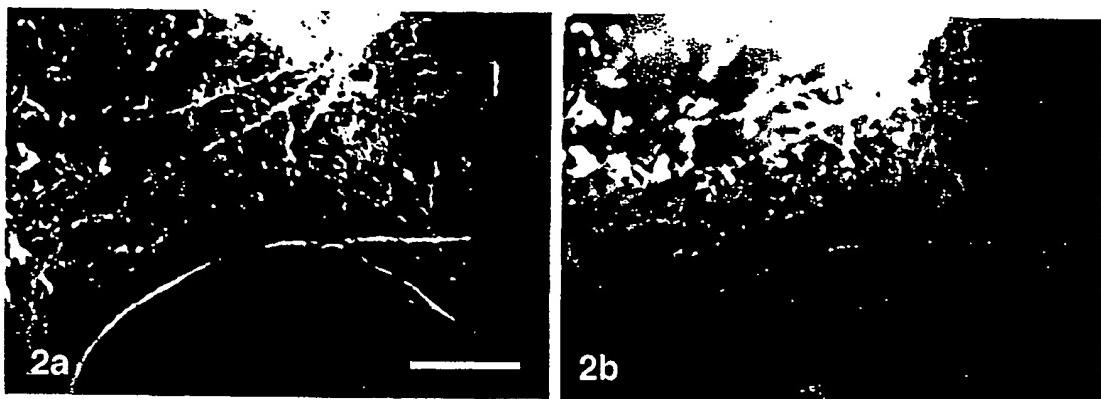


FIG. 20

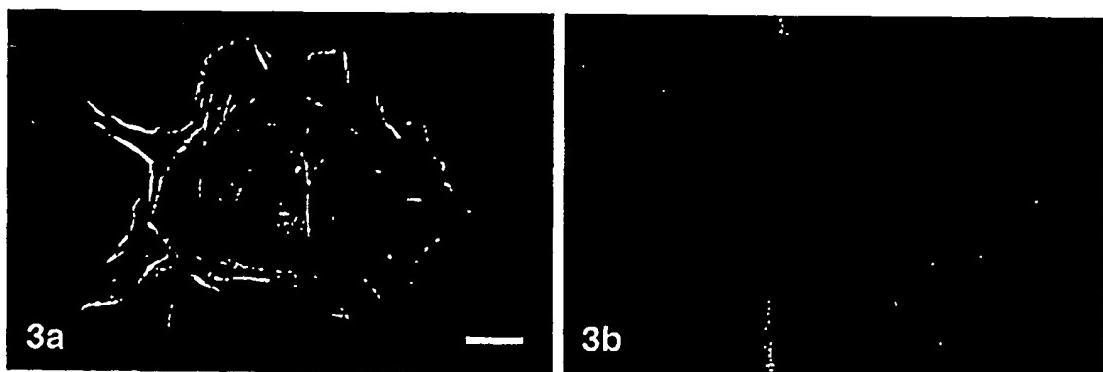


FIG. 21 BEST AVAILABLE COPY

24/26

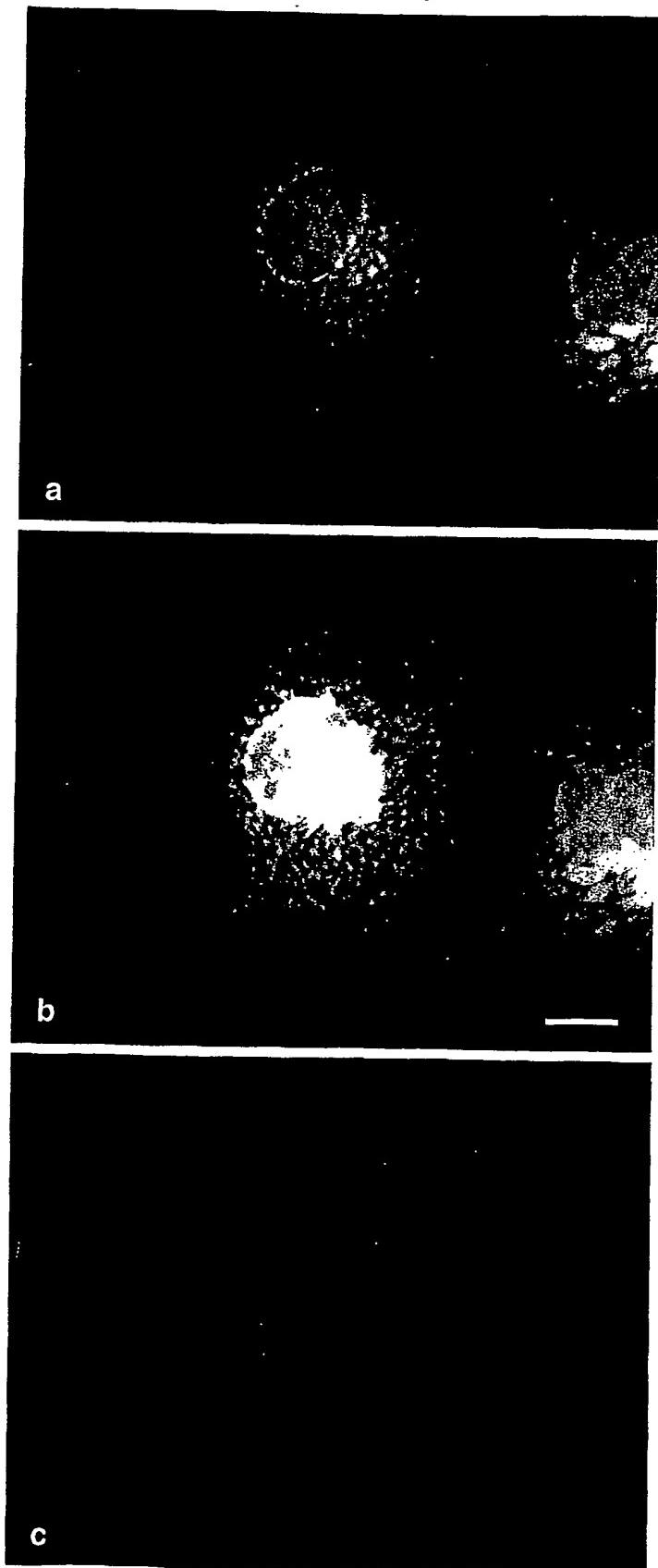


FIG. 22

BEST AVAILABLE COPY

25/26

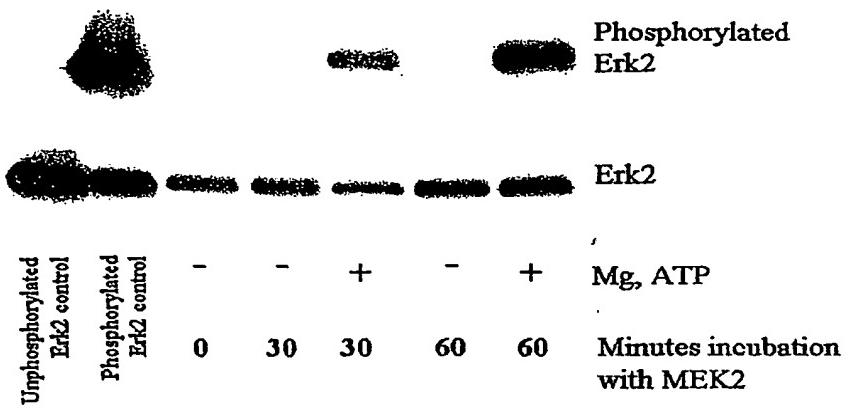


FIG. 23

BEST AVAILABLE COPY

26/26

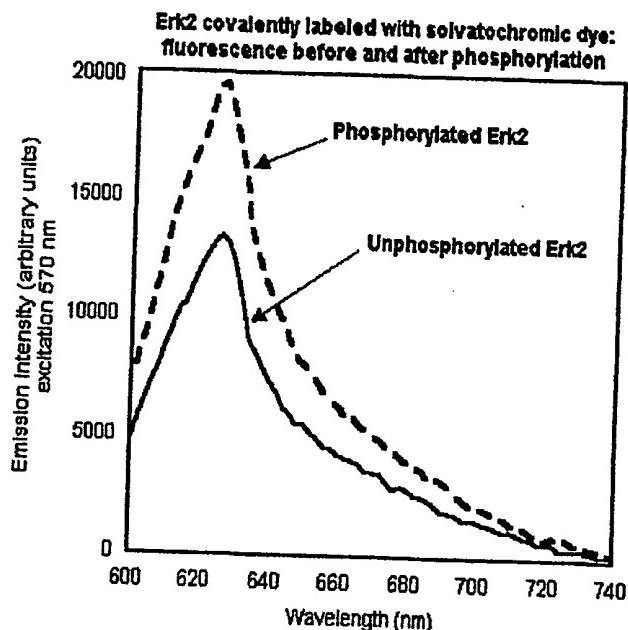


FIG. 24

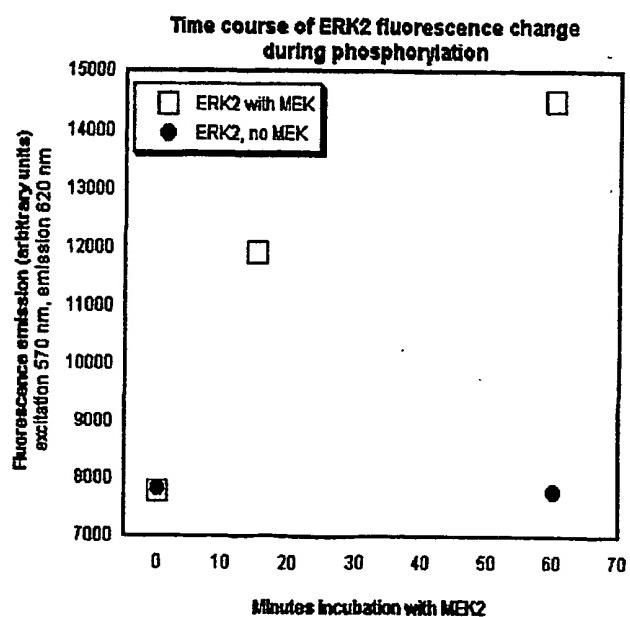


FIG. 25

SEQUENCE LISTING

<110> The Scripps Research Institute et al.

5

<120> Labeled peptides, proteins and antibodies and processes and
intermediates useful for their preparation

<130> 1361.007WO1

10

<150> US 09/839,577

<151> 2001-04-20

<160> 26

15

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 44

20<212> PRT

<213> Homo sapiens

<400> 1

Lys Lys Lys Glu Lys Glu Arg Pro Glu Ile Ser Leu Pro Ser Asp Phe

25 1

5

10

15

Glu His Thr Ile His Val Gly Phe Asp Ala Cys Thr Gly Glu Phe Thr

20

25

30

Gly Met Pro Glu Gln Trp Ala Arg Leu Leu Gln Thr

35

40

30

<210> 2

<211> 16

<212> PRT

<213> Artificial Sequence

35

<220>

<223> A synthetic peptide.

<400> 2

40Ala Lys Ala Ala Arg Ala Ala Ala Lys Ala Ala Arg Ala Cys Ala

1

5

10

15

<210> 3
<211> 6
<212> PRT
5<213> Artificial Sequence

<220>
<223> A synthetic peptide.

10<221> SITE
<222> 3
<223> Xaa = SAOD =
alpha-Boc-beta [N- (2-Chlorobenzylloxycarbonyl) -N-Methylaminoxy
Acetyl] -alpha,beta-Diaminopropionic Acid [Boc-2-Cl-Z-(SA)Dapa-OH].

15
<221> SITE
<222> 6
<223> Xaa = MPAL = The C-terminal mercaptopropionyl-leucine group
generated by cleavage of a peptide from TAMPAL resin.

20
<400> 3
Leu Tyr Xaa Ala Gly Xaa
1 5

25<210> 4
<211> 5
<212> PRT
<213> Artificial Sequence

30<220>
<223> A synthetic peptide.

<400> 4
Cys Arg Ala Asn Lys
35 1 5

<210> 5
<211> 10
<212> PRT
40<213> Artificial Sequence

<220>

<223> A synthetic peptide.

<221> SITE

5<222> 3

<223> Xaa = SAOD =

alpha-Boc-beta[N-(2-Chlorobenzylloxycarbonyl)-N-Methylaminoxy
Acetyl]-alpha,beta-Diaminopropionic Acid [Boc-2-Cl-Z-(SA)Dapa-OH]

10<400> 5

Leu Tyr Xaa Ala Gly Cys Arg Ala Asn Lys

1 5 10

<210> 6

15<211> 33

<212> PRT

<213> Artificial Sequence

<220>

20<223> A synthetic peptide.

<400> 6

Cys Glu Tyr Arg Ile Asp Arg Val Arg Leu Phe Val Asp Lys Leu Asp

1 5 10 15

25Asn Ile Ala Gln Val Pro Arg Val Gly Ala Ala His His His His

20 25 30

His

30<210> 7

<211> 33

<212> PRT

<213> Artificial Sequence

35<220>

<223> A synthetic peptide.

<400> 7

Cys Glu Tyr Arg Ile Asp Arg Val Arg Leu Phe Val Asp Lys Leu Asp

40 1 5 10 15

4

Asn Ile Ala Gln Val Pro Arg Val Gly Ala Ala His His His His His
20 25 30
His

5

<210> 8
<211> 28
<212> PRT
<213> Artificial Sequence

10

<220>
<223> A synthetic peptide.

<221> SITE

15<222> 1

<223> Xaa = SAOD =
alpha-Boc-beta[N-(2-Chlorobenzylloxycarbonyl)-N-Methylaminoxy
Acetyl]-alpha,beta-Diaminopropionic Acid [Boc-2-Cl-Z-(SA)Dapa-OH]

20<221> SITE

<222> 28

<223> Xaa = MPAL = The C-terminal mercaptopropionyl-leucine group
generated by cleavage of a peptide from TAMPAL resin.

25<400> 8

Xaa Lys Lys Lys Glu Lys Glu Arg Pro Glu Ile Ser Leu Pro Ser Asp
1 5 10 15
Phe Glu His Thr Ile His Val Gly Phe Asp Ala Xaa
20 25

30

<210> 9
<211> 18
<212> PRT
<213> Homo sapiens

35

<400> 9
Cys Thr Gly Glu Phe Thr Gly Met Pro Glu Gln Trp Ala Arg Leu Leu
1 5 10 15
Gln Thr

40

5

<210> 10
<211> 10
<212> PRT
<213> Artificial Sequence

5

<220>
<223> A synthetic peptide.

<221> SITE
10<222> 3
<223> Xaa = SAOD =
alpha-Boc-beta[N-(2-Chlorobenzylloxycarbonyl)-N-Methylaminoxy
Acetyl]-alpha,beta-Diaminopropionic Acid [Boc-2-Cl-Z-(SA)Dapa-OH]

15<400> 10
Leu Tyr Xaa Ala Gly Cys Arg Ala Asn Lys
1 5 10

<210> 11
20<211> 45
<212> PRT
<213> Artificial Sequence

<220>
25<223> A synthetic peptide.

<221> SITE
<222> 1
<223> Xaa = SAOD =
30alpha-Boc-beta[N-(2-Chlorobenzylloxycarbonyl)-N-Methylaminoxy
Acetyl]-alpha,beta-Diaminopropionic Acid [Boc-2-Cl-Z-(SA)Dapa-OH]

<400> 11
Xaa Lys Lys Lys Glu Lys Glu Arg Pro Glu Ile Ser Leu Pro Ser Asp
35 1 5 10 15
Phe Glu His Thr Ile His Val Gly Phe Asp Ala Cys Thr Gly Glu Phe
20 25 30
Thr Gly Met Pro Glu Gln Trp Ala Arg Leu Leu Gln Thr
35 40 45

<210> 12
<211> 38
<212> PRT
<213> Artificial Sequence

5

<220>
<223> A synthetic peptide.

<221> SITE
10<222> 3
<223> Xaa = SAOD =
alpha-Boc-beta[N-(2-Chlorobenzylloxycarbonyl)-N-Methylaminoxy
Acetyl]-alpha,beta-Diaminopropionic Acid [Boc-2-Cl-Z-(SA)Dapa-OH]

15<400> 12
Leu Tyr Xaa Ala Gly Cys Glu Tyr Arg Ile Asp Arg Val Arg Leu Phe
1 5 10 15
Val Asp Lys Leu Asp Asn Ile Ala Gln Val Pro Arg Val Gly Ala Ala
20 25 30
20His His His His His
35

<210> 13
<211> 120
25<212> PRT
<213> Homo sapiens

<400> 13
Asp Ile Gln Asn Pro Asp Ile Thr Ser Ser Arg Tyr Arg Gly Leu Pro
30 1 5 10 15
Ala Pro Gly Pro Ser Pro Ala Asp Lys Lys Arg Ser Gly Lys Lys Lys
20 25 30
Ile Ser Lys Ala Asp Ile Gly Ala Pro Ser Gly Phe Lys His Val Ser
35 40 45
35His Val Gly Trp Asp Pro Gln Asn Gly Phe Asp Val Asn Asn Leu Asp
50 55 60
Pro Asp Leu Arg Ser Leu Phe Ser Arg Ala Gly Ile Ser Glu Ala Gln
65 70 75 80
Leu Thr Asp Ala Glu Thr Ser Lys Leu Ile Tyr Asp Phe Ile Glu Asp
40 85 90 95

Gln Gly Gly Leu Glu Ala Val Arg Gln Glu Met Arg Arg Gln Glu Pro

100 105

110

Leu Pro Pro Pro Pro Pro Ser

115 120

5

<210> 14

<211> 502

<212> PRT

10<213> Homo sapiens

<400> 14

Met Ser Gly Gly Pro Met Gly Gly Arg Pro Gly Gly Arg Gly Ala Pro

1 5 10 15

15 Ala Val Gln Gln Asn Ile Pro Ser Thr Leu Leu Gln Asp His Glu Asn

20 25 30

Gln Arg Leu Phe Glu Met Leu Gly Arg Lys Cys Leu Thr Leu Ala Thr

35 40 45

Ala Val Val Gln Leu Tyr Leu Ala Leu Pro Pro Gly Ala Glu His Trp

20 50 55 60

Thr Lys Glu His Cys Gly Ala Val Cys Phe Val Lys Asp Asn Pro Gln

65 70 75 80

Lys Ser Tyr Phe Ile Arg Leu Tyr Gly Leu Gln Ala Gly Arg Leu Leu

85 90 95

25 Trp Glu Gln Glu Leu Tyr Ser Gln Leu Val Tyr Ser Thr Pro Thr Pro

100 105 110

Phe Phe His Thr Phe Ala Gly Asp Asp Cys Gln Ala Gly Leu Asn Phe

115 120 125

Ala Asp Glu Asp Glu Ala Gln Ala Phe Arg Ala Leu Val Gln Glu Lys

30 130 135 140

Ile Gln Lys Arg Asn Gln Arg Gln Ser Gly Asp Arg Arg Gln Leu Pro

145 150 155 160

Pro Pro Pro Thr Pro Ala Asn Glu Glu Arg Arg Gly Gly Leu Pro Pro

165 170 175

35 Leu Pro Leu His Pro Gly Gly Asp Gln Gly Gly Pro Pro Val Gly Pro

180 185 190

Leu Ser Leu Gly Leu Ala Thr Val Asp Ile Gln Asn Pro Asp Ile Thr

195 200 205

Ser Ser Arg Tyr Arg Gly Leu Pro Ala Pro Gly Pro Ser Pro Ala Asp

40 210 215 220

Lys Lys Arg Ser Gly Lys Lys Lys Ile Ser Lys Ala Asp Ile Gly Ala
225 230 235 240
Pro Ser Gly Phe Lys His Val Ser His Val Gly Trp Asp Pro Gln Asn
245 250 255
5Gly Phe Asp Val Asn Asn Leu Asp Pro Asp Leu Arg Ser Leu Phe Ser
260 265 270
Arg Ala Gly Ile Ser Glu Ala Gln Leu Thr Asp Ala Glu Thr Ser Lys
275 280 285
Leu Ile Tyr Asp Phe Ile Glu Asp Gln Gly Leu Glu Ala Val Arg
10 290 295 300
Gln Glu Met Arg Arg Gln Glu Pro Leu Pro Pro Pro Pro Pro Ser
305 310 315 320
Arg Gly Gly Asn Gln Leu Pro Arg Pro Pro Ile Val Gly Gly Asn Lys
325 330 335
15Gly Arg Ser Gly Pro Leu Pro Pro Val Pro Leu Gly Ile Ala Pro Pro
340 345 350
Pro Pro Thr Pro Arg Gly Pro Pro Pro Pro Gly Arg Gly Pro Pro
355 360 365
Pro Pro Pro Pro Pro Ala Thr Gly Arg Ser Gly Pro Leu Pro Pro Pro
20 370 375 380
Pro Pro Gly Ala Gly Gly Pro Pro Met Pro Pro Pro Pro Pro Pro Pro
385 390 395 400
Pro Pro Pro Pro Ser Ser Gly Asn Gly Pro Ala Pro Pro Pro Leu Pro
405 410 415
25Pro Ala Leu Val Pro Ala Gly Gly Leu Ala Pro Gly Gly Arg Gly
420 425 430
Ala Leu Leu Asp Gln Ile Arg Gln Gly Ile Gln Leu Asn Lys Thr Pro
435 440 445
Gly Ala Pro Glu Ser Ser Ala Leu Gln Pro Pro Pro Gln Ser Ser Glu
30 450 455 460
Gly Leu Val Gly Ala Leu Met His Val Met Gln Lys Arg Ser Arg Ala
465 470 475 480
Ile His Ser Ser Asp Glu Gly Glu Asp Gln Ala Gly Asp Glu Asp Glu
485 490 495
35Asp Asp Glu Trp Asp Asp
500

<210> 15

<211> 34

40<212> PRT

<213> Artificial Sequence

<220>

<223> A synthetic peptide.

5

<400> 15

Cys Glu Met Ala Gln Leu Glu Lys Glu Val Gln Ala Leu Glu Ser Glu
1 5 10 15
Val Ala Ser Leu Glu Lys Glu Val Gln Ala Leu Glu Lys Glu Val Ala
10 20 25 30
Gln Arg

<210> 16

15<211> 33

<212> PRT

<213> Artificial Sequence

<220>

20<223> A synthetic peptide.

<400> 16

Lys Met Ala Gln Leu Lys Lys Lys Val Gln Ala Leu Lys Ser Lys Val
1 5 10 15
25Ala Ser Leu Lys Lys Lys Val Gln Ala Leu Lys Lys Lys Val Ala Gln
20 25 30
Arg

30<210> 17

<211> 12

<212> PRT

<213> Artificial Sequence

35<220>

<223> A synthetic peptide.

<400> 17

Tyr Gly Arg Lys Lys Arg Arg Asn Arg Arg Arg Pro
40 1 5 10

10

<210> 18
<211> 39
<212> DNA
<213> Artificial Sequence

5

<220>
<223> A synthetic oligonucleotide.

<400> 18
10gagctgtaca agggaggcat gaagatggcc cagctgaag

39

<210> 19
<211> 34
<212> DNA
15<213> Artificial Sequence

<220>
<223> A synthetic oligonucleotide.

20<400> 19
gtcgcggccg ccttatacgct gggccacctt cttc

34

<210> 20
<211> 37
25<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic oligonucleotide.

30

<400> 20
gactggatcc gagtccatga agaagatggc tcctgcc

37

<210> 21
35<211> 66
<212> DNA
<213> Artificial Sequence

<220>
40<223> A synthetic oligonucleotide.

11

<400> 21
ctgaactttc ttcttgagct gggccatctt acttccgccca ccgttgtttt tggggtttgg 60
cttagg 66

5<210> 22
<211> 63
<212> DNA
<213> Artificial Sequence

10<220>
<223> A synthetic oligonucleotide.

<400> 22
caccttttc ttcaagctcg ccactttgct cttgagcgcc tgaactttct tcttgagctg 60
15gac 63

<210> 23
<211> 66
<212> DNA
20<213> Artificial Sequence

<220>
<223> A synthetic oligonucleotide.

25<400> 23
ggcctcgagt caacgttgag cgactttctt ttttaaggct tgcacctttt tcttcaagct 60
cgccac 66

<210> 24
30<211> 35
<212> DNA
<213> Artificial Sequence

<220>
35<223> A synthetic oligonucleotide.

<400> 24
ctcgatcgct cgagtcaacg ttgagcgact ttctt 35

40<210> 25

12

<211> 20
<212> DNA
<213> Artificial Sequence

5<220>
<223> A synthetic oligonucleotide.

<400> 25
ggggatcctc tcagacgccc

10
<210> 26
<211> 38
<212> DNA
<213> Artificial Sequence

15
<220>
<223> A synthetic oligonucleotide.

<400> 26
20catgccatgg ctccggatcc accagctgtg agagacac

20

38